The Effects of Pre-Task Planning on Fluency, Complexity and Accuracy in L2 Oral Narrative Tasks

MOK, Joyce Mee Luen

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Philosophy

in

English (Applied English Linguistics)

August 2007

Thesis / Assessment Committee

Professor Peter Skehan (Chair)

Professor Jette Hansen Edwards (Thesis Supervisor)

Professor Gwendolyn Gong (Committee Member)

Professor Cynthia Lee (External Examiner)

ABSTRACT

This thesis explores the effects of pre-task planning on L2 learners' language production. Research to date supports the claim that fluency and complexity are positively influenced by pre-task planning. However, mixed results are found for accuracy. This study looks for possible explanations by utilizing alternative prompts for narrative tasks – video snippets – and comparing the results with narrative tasks using picture sequences. Eight intermediate learners and eight advanced learners of English performed a series of tasks, followed by post-task questionnaires.

Results conform to previous research, showing that pre-task planning does enhance fluency and complexity, with the exception of accuracy. Learners expressed more enjoyment and less difficulty with tasks involving video snippets, though video tasks are not particularly indicative of improvement in language production.

Regardless of their proficiency level, learners expressed planning strategies which are consistent with categories identified by Ortega (2004). These categories include rehearsal and retrieval operations. These findings help to further our understanding of the interrelationship between planning and L2 oral output, as they indicate alternatives for the traditional use of prompts in narrative tasks, and also the consideration of including the learners' perspective in research in this predominantly quantitative field. Further implications for language pedagogy are also discussed.

導論

這個研究探討任務前計劃如何影響香港學生的英語說話能力。迄今的研究指出說話的流暢度及語句的複雜性都受到事前計劃的正面影響。換言之，任務前計劃能夠令語言學習者的說話更加流暢及多元化。不過有研究指出任務前計劃對於學習者說話的準確性，不一定有正面的影響。傳統的敍述性任務均使用圖畫作爲給予學習者的提示，然而這個研究會以短片作爲另類的提示，誘導學習者敍述故事。另外，給予學習者的任務亦會分爲具備事前計劃及沒有事前計劃。

這個研究探討香港中文大學 16 位英文系主修生如何以英語敍述故事。他們將會分成兩個組別：高等程度的學習者 (於香港高級程度會考英語應用科目取得 A 或 B 的成績)以及中等程度的學習者 (於香港高級程度會考英語應用科目取得 C 或 D 的成績)。兩個組別分別進行一連串的任務，然後再完成問卷調查，問卷調查的內容圍繞他們對自己所進行的任務有何意見。

研究結果顯示英語學習者在具備任務前計劃的情況下，的確能夠較爲流暢及多元化地敍述一個故事，不過任務前計劃對於說話的準確性並沒有顯著的幫助。有關短片的應用，大多數的學習者均表示更加享受任務的過程；再者，學習者的英語程度對他們的故事敍述並沒有太大的影響。總括而言，這個研究給予未來對於這個範疇的人不少的新方向。

ACKNOELDGEMENTS

I would like to express my deepest gratitude to my supervisor, Professor Jette Hansen Edwards, for her encouragement and support. Her guidance is crucial to the successful completion of this thesis.

Next I would like to thank Professor Peter Skehan for his continuous support throughout my studies at this university. He has given me some insightful suggestions for the topic of my research.

I would also like to thank Professor Gwendolyn Gong for her advice on my thesis. Her useful comments have further helped me refine my thesis.

Last but not the least, I would like to thank my participants for their enthusiasm in my study. Without their help this study would neither be successful nor possible.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background to this study

"All spoken and written language use, even that which appears effortless and automatic, involves planning" (Ellis, 2005, p. 3). The importance of planning in the execution of a language act is ascertained by the amount of literature that has focused on the influence of task planning on L2 learners' language production (Bygate, 1996; Foster & Skehan, 1996, 1997; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003). Two types of planning are identified by Ellis (2005), namely pre-task planning and on-line planning. Generally speaking, learners who are provided with the opportunity to plan (i.e. provided with planning time) before task performance undergo pre-task planning, while learners who monitor their language production when speaking undergo on-line planning. Because of the overwhelming amount of literature on pre-task planning, this will also be the focus of this study.

Studies on pre-task planning are based upon the theory of information processing, which states that humans possess a limited working memory which can only process a certain amount of information at a time (Anderson, 1995). Skehan's (1998) framework of tasks takes into account the limited attentional resources that language learners possess, and claims that various task characteristics, including pre-task

planning, can be manipulated to influence the three aspects of speaking, namely fluency, complexity and accuracy. Because pre-task planning can lessen the cognitive load of learners' working memory while they process the task, Skehan (1998) claims that when given the opportunity to plan before the task, learners are able to produce more fluent and complex language.

The positive effects of pre-task planning on language production are supported by a large amount of research (Bygate, 1996; Foster & Skehan, 1996, 1997; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003). Bygate (1996) explored task familiarity and found that as learners become more familiar with the task, their syntactic complexity also increases. Foster and Skehan (1996) found that pre-task planning does contribute to fluency under conditions where learners are given some guidance on how to plan their speech. Mehnert (1998) studied different lengths of planning time (no planning, 1-minute, 5-minute and 10-minute) and found that fluency increases with the amount of planning time, whereas complexity is greatest with 10-minute planners. Yuan and Ellis (2003) also found that pre-task planning enhances fluency and complexity.

In general, most studies (Bygate, 1996; Foster & Skehan, 1996, 1997; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003) on pre-task planning have come to the conclusion that pre-task planning has positive effects on the fluency and complexity of L2 learners' speech. Not all studies, however, have found the same for accuracy. Skehan and Foster (1997), Mehnert (1998)

and Tavakoli and Skehan (2005) are some of the studies which have found that accuracy is enhanced with pre-task planning under certain conditions, such as the task type, length of planning time and whether an inherent structure is present in the task, and so on. Other studies (Wigglesworth, 1997; Yuan & Ellis, 2003) found no positive effects of pre-task planning on accuracy. Thus unlike fluency and complexity, the effects of pre-task planning on accuracy is still a controversial issue. This leads to questions about what other task characteristics might influence fluency, complexity and accuracy, and whether pre-task planning leads to a positive effect in accuracy.

It can be generalized from the above that no pre-task planning study has found the increase of all three aspects of speaking under the same conditions, which is consistent with the information-processing approach, which states that humans possess a limited processing capacity, which does not allow us to attend fully to all aspects of a task (Anderson, 1995). Using this as the rationale, Skehan's (1998) framework of tasks suggests that there is a trade-off effect between fluency, accuracy and complexity when learners are processing speech. He argues that mutual tension exists between accuracy and complexity, which compete with each other for the limited attentional resources of an L2 learner, resulting in either an enhancement of accuracy or complexity in a task, to the detriment of the other. This claim is supported by many studies (Foster & Skehan, 1996; Mehnert, 1998; Skehan & Foster, 1997; Tavakoli & Skehan, 2005), indicating that under the same task, pre-task planning only has strong positive effect on either complexity or accuracy, but not both at the same

time.

Interestingly, studies that investigated both on-line planning and pre-task planning have found that the trade-off effect exists between accuracy and fluency, instead of between accuracy and complexity (Wendel, 1997, as cited in Yuan & Ellis, 2003; Yuan & Ellis, 2003). Wendel (1997) suggested that pre-task planning indeed has positive effects on fluency, but that it is *online planning* that increases the accuracy of learners' speech. Yuan and Ellis (2003) also came to the same conclusion. This raises the question of whether the trade-off effect influences accuracy and fluency, or accuracy and complexity or whether it affects both aspects under different task conditions.

As this brief introduction to research on task planning indicates, debatable issues still remain as to the task characteristics that influence fluency, complexity and accuracy. This thesis explores these issues by examining the influence of different task characteristics on the oral performance of L2 learners. Pre-task planning is one of the independent variables of this study, which will make use of narrative tasks as the main testing tools to elicit learners' speech. This is to ensure comparability with previous research.

On the other hand, the studies mentioned above also draw our attention to the traditional use of prompts in pre-task planning research, which shows that picture sequences are the major testing material in narrative tasks (Ortega, 1999; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Wigglesworth, 1997; Yuan & Ellis, 2003).

Undoubtedly picture sequences do provide a context for participants to tell a story, and are justified in terms of construct validity, reliability, and the authenticity of tests. Despite this, a limited number of studies (cf. Bygate, 1996; Gass & Mackey, 1999; Skehan & Foster, 1999; Wendel, 1997, as cited in Yuan & Ellis, 2003) make use of videos as the prompt for narrative tasks. Although none of these studies have actually specified the benefits of video prompts on learners' speech, the suggestion of using video snippets as prompts for narrative tasks provide an interesting alternative to traditional picture sequences. Studies to date on pre-task planning have never compared the use of picture prompts and video prompts in narrative tasks, or whether there is a difference in the elicited language performance on these two prompts. This raises the interesting question of whether learners produce the same level of fluency, accuracy and complexity with video prompts, or whether videos, which provide so much more audio-visual stimulation, will enhance certain aspects of their performance more than pictures.

Another important issue concerns the learners' perception of tasks. A limited number of studies have explored learner strategies during pre-task planning (Ortega, 1999, 2005) and learners' perceptions of task difficulty and the usefulness of pre-task planning (Elder et al., 2002; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003). Either post-task questionnaires or retrospective semi-structured interviews were used in these studies. Ortega (1999, 2005) found that language expertise is a factor in differences in task performance although it does not lead to much difference in

strategy use during pre-task planning. Elder, Iwashita, and McNamara (2002) attempted to find a correlation between learners' attitude and task performance, but the weak correlations showed that the perception of task difficulty is a multi-dimensional phenomenon which involves complex and unstable interactions between different task features and test-taker attributes.

Other than those mentioned above, studies on planning to date have provided very little information on what learners did while they were performing the task, which is clearly important in order to understand where pre-task planning benefits come from. This raises the following questions: What do learners do while they plan? What difficulties do they face? What different strategies will learners of different proficiency levels use? These issues are explored in order to provide a different dimension to this basically quantitative field. The way in which learner's perspective affects their task performance is not a main issue in this study.

1.2 Purpose of this study

Drawing from the literature on pre-task planning, this thesis aims to answer the following questions:

1. What effect does pre-task planning, in addition to the use of picture prompts and video prompts, have on the *fluency* of L2 learners' production in an oral narrative task?
2. What effect does pre-task planning, in addition to the use of picture prompts and

video prompts, have on the *complexity* of L2 learners' production in an oral narrative task?

3. What effect does pre-task planning, in addition to the use of picture prompts and video prompts, have on the *accuracy* of L2 learners' production in an oral narrative task?
4. What are learners' perceptions of the usefulness of pre-task planning time, task enjoyment or task difficulty? What do learners do when they plan?

This study is operationalized according to Skehan's (1998) framework of tasks, which also forms the basis of the research questions above. In laboratory settings, all learners performed a series of narrative tasks involving picture sequences and video snippets under different planning conditions. In addition, retrospective questionnaires on task difficulty were used to elicit information on learners' attitude and perspective of their own task performance. This triangulation of data sources ensures that the issue of pre-task planning effects will be explored in depth in this study.

1.3 Organization of this thesis

The thesis is divided into six chapters. This introduction to pre-task planning literature aims to give readers a general idea of the key concepts and issues in this field. The following chapters provide readers with the solid content of this study.

Chapter 2 – Literature Review: This chapter starts with providing the definitions of key concepts and terms, such as the information-processing theory, Levelt's (1989)

model of speech, and Skehan's (1998) cognitive approach to tasks. Alternative approaches are also discussed. A brief overview of task-based literature, both quantitative and qualitative, together with their findings on the effects of pre-task planning on L2 learners' speech, are also discussed.

Chapter 3 – Method: This chapter discusses a number of methodological issues and presents the major theoretical and analytical framework employed in this study. Details of operationalization are also provided here. Apart from accounting for the source and selection of participants and task material, other issues discussed include the justification for and limitations of the task design and data collection procedures. Last of all, this chapter explains the procedures of analyzing participants' speech and collecting the questionnaire data.

Chapter 4 – Results: This chapter reports and discusses the findings of the present investigation. The fluency, accuracy and complexity patterns of participants' speech data is presented descriptively. Information collected from the task difficulty questionnaire is also discussed. The interpretation of data is reserved for the following chapter, which is focused on explaining the findings with reference to the existing literature.

Chapter 5 – Discussion: This chapter offers explanations and justifications for the results presented in the previous chapter. The research questions raised in Chapter 2 (Literature Review) are also answered, based on the findings described in Chapter 4 (Results). Relationships are drawn between the existing literature and the findings of

this study.

Chapter 6 – Conclusion: This chapter sums up the study and discusses its significance to the growing amount of research as a whole. Suggestions for future research are provided in discussing the limitations of the present study, representing an effort to contribute to the existing literature which explores the effects of pre-task planning on L2 learners' language production under testing conditions. Hopefully both the quantitative and qualitative perspectives of this study will add a fresh dimension to this traditionally researcher-oriented field.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

Over the past decade, a number of studies have investigated the effects of pre-task planning on language production (cf. Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita, McNamara, & Elder., 2001; Ortega, 1999, 2005; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003). Specifically, these studies have explored the effect of pre-task planning on learners' language production in oral tasks, and how their task performance is affected by various task characteristics.

This section first introduces some of the basic concepts which form the rationale of literature on pre-task planning, including information-processing theory, Levelt's (1989) model of speech and different cognitive approach of tasks (i.e. Skehan's (1998) framework of tasks and Robinson's (2001) alternative approach). In the second part of this section, the different types of planning (pre-task planning and on-line planning) are discussed, followed by different planning conditions that were explored in previous research including guided planning vs. unguided planning (Foster & Skehan, 1996), group planning vs. individual planning (Foster & Skehan, 1999), personal tasks vs. narrative tasks vs. decision-making tasks (Skehan & Foster, 1997), group planning vs. solitary planning (Foster & Skehan, 1999), different lengths of planning

time (Mehnert, 1998), structured tasks vs. unstructured tasks (Tavakoli & Skehan, 2005) and language expertise (Ortega, 1999, 2005). Finally, how learners' language production is affected by these different planning conditions, and how this relates to the basis of the present study, are discussed.

2.2 Cognitive Approaches to Task Performance

Skehan's (1998) framework of tasks and Robinson's (2001) alternative view to task performance drew on the information-processing theory. Before discussing both frameworks, the information-processing theory is first discussed, followed by Baddeley's (Baddeley & Hitch, 1974; Baddeley & Logie, 1999) model of working memory and Levelt's (1989) model of speech production. These two models are most commonly cited in task planning literature (Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Ortega, 1999, 2005; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003).

2.2.1 The Information-Processing Theory

Literature that has focused on pre-task planning more or less drew on information-processing theory as the justification for their studies (Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Ortega, 1999, 2005; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003). This theory is driven by a number of general principles. The first principle that

informed this theory is the limited capacity of the working memory, which is only able to process limited amount of information. The limitation of the working memory is likely to cause bottlenecks and lead to language users prioritizing one aspect of language production over another. The implications of this to learners' language production are discussed later on. The second principle is the control mechanism that language users access when they are not familiar with a task. This mechanism draws on explicit stored knowledge when language users are faced with a new task for which they do not possess proceduralized knowledge, thus using up processing power and taxing the working memory of language users. Again this points at the cognitive load the working memory faces during language processing.

In short, the limitation in working memory capacity *and* the unfamiliarity of a task increase the cognitive burden on the working memory. But how does this affect language production? In order to understand how the working memory runs in the human brain, Baddeley's (Baddeley & Hitch, 1974; Baddeley & Logie, 1999) model of working memory, one of the most commonly cited models in task planning literature (cf. Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Ortega, 1999, 2005; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003), must first be explained.

Baddeley's (Baddeley & Hitch, 1974; Baddeley & Logie, 1999) model of working memory identifies three components of the working memory, including the central executive, the phonological loop and the visual spatial sketchpad. The central

executive system and the phonological loop are relevant to understanding the role for planning in task performance. The relationship between the working memory and long-term memory is governed by the central executive system, which is limited in capacity. Thus language users have to make decisions as to where to allocate their available processing space. The phonological loop is made up of two components, namely the phonological store and a mechanism allowing for articulatory rehearsal. The latter of these components enables linguistic input in the working memory to be sustained, thus planning is likely to draw extensively on this component, as it allows language users to maintain one set of material while drawing on another set to modify it (Ellis, 2005). In short, the phonological loop plays a central role in monitoring. In terms of the information-processing approach and Baddeley's model of working memory, task planning for language users represents a way to help overcome the limitations of attentional resources in the working memory.

The limitation of the working memory capacity prevents the language learner from conveying the meaning and conforming to grammatical rules during task performance at the same time, thus Van Patten (1990) suggested that to enhance the grammatical accuracy of learners' language production in a task, language tasks should bring their attention to target grammatical forms. In other words, language tasks should include a "focus-on-form." In the context of language pedagogy, focus-on-form refers to "attempts to intervene in the process of acquisition by inducing learners to pay attention to linguistic form while they are primarily

concerned with decoding or encoding message content" (Ellis, 2005, p. 9). With a limited working memory capacity, language users may find it difficult to attend to both the meaning of the task and its linguistic form at the same time, hence they usually give priority to decoding the meaning of the task, whereas non-salient linguistic features are most likely to be overlooked. To compensate for this, providing learners with the opportunity to plan (i.e. extra time to prepare for their task performance) offers learners with a means to achieve a balance between message conveyance and linguistic accuracy at the same time.

To understand how the information-processing theory and Baddeley's (Baddeley & Hitch, 1974; Baddeley & Logie, 1999) working memory model apply to learners' language production, the following section will introduce the Levelt's (1989) model of speech production, which is followed by the description of different cognitive approaches towards task performance, and the elaboration of the three key aspects of speaking, which are fluency, accuracy and complexity.

2.2.2 Levelt's Model of Speech Production

Where studies (Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Ortega, 1999, 2005; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003) of task planning are concerned, Levelt's (1989) model of speech production is drawn upon as the rationale for learners' speech. This model identifies three autonomous processing stages: 1) conceptualization, 2) formulation,

and 3) articulation.

Conceptualization is a stage where the speaker first of all decides upon the communicative goal, after which she/he develops this communicative goal into a series of sub-goals, and identifies a speech act for each sub-goal that would achieve the intended effect. Finally, she/he retrieves the necessary information to realize each of the subgoals. By determining various issues which include the topic and how it would interest the listener, the speaker organizes the information, which then becomes a preverbal but non-linguistic message. This message is then forwarded to the formulator.

The second stage of speech production, formulation, retrieves lexical items from the speaker's mental lexicon and establishes language representations of the preverbal messages. This serves to prompt the syntactic building procedure required for grammatical encoding, resulting in a "surface structure" (Levelt, 1989, p. 11) which is processed and converted into an articulatory plan. Levelt (1989) calls this "internal speech" (p. 12), which is transferred to the articulator.

The last stage of speech production, articulation, is involved with the retrieval of chunks of internal speech. Ultimately this leads to the production of overt speech. During speech production, a self-monitoring process regulates and inspects these three stages under operation.

These three stages of speech production show that the latter two processes, formulation and articulation, relates to the building up of the lexical, syntactic and

grammatical structures of an utterance, and also the constant inspection of these structures to ensure that the desired effect is achieved. As this is a matter of language expertise and familiarity with the target language, native speakers of the target language are able to carry out the processes automatically, whereas similar speech production may be demanding on the working memory of language learners, as they must activate and execute their linguistic knowledge through controlled processing (as opposed to automatic processing), which means that learners are conscious of the retrieval and building up of linguistic resources. This shows that language learners, especially those who are limited in proficiency, are most likely to experience problems in achieving linguistic fluency, syntactic and/or lexical complexity, or grammatical accuracy.

2.2.3 Cognitive Models of Task-Based Performance

While Levelt's (1989) model of speech production forms the basic concept to speech processing, it is Skehan's (1998) framework of tasks which concretizes the key aspects of language and how these aspects interact with each other. Skehan's framework is widely adopted by a number of studies on task planning (Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Skehan & Foster, 1997; Ortega, 1999, 2005; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003). He distinguishes three aspects of production -- fluency, accuracy and complexity -- and suggests that learners vary in the extent to which they emphasize each aspect. The

definitions of these three measures are as follows (Skehan, 1996, 1998):

1. *Fluency* is "learners' capacity to produce language in real time without undue pausing or hesitation" (Skehan, 1996, p. 22). Fluency can be affected by tasks which are high in cognitive demand, and cause the speaker to experience difficulty in retrieving a lexical or syntactic item during online processing. Skehan (1998) distinguishes between time-creating dysfluency (meaningless phrases such as "you know", "actually"), breakdown dysfluency ("uh", "um", silence at undue places such as in the middle of a clause, but not at clause boundaries) and repair dysfluency (reformulation, false starts, repetition, replacements) (Skehan, 1996).

2. *Complexity* is "the elaboration or ambition of language that is produced. It involves lexical content and syntactic usage" (Skehan, 1996, p. 22). It is usually categorized into three groups by researchers (cf. Skehan & Foster, 1997, 1999; Yuan & Ellis, 2003) including *semantic complexity* (the less the words are used to convey an idea, the more complex the sentence), *syntactic complexity* (the more subordinate clauses there are in a speech unit, the more complex the sentence), and *lexical variety* (the more difficult it is to retrieve certain words for expression, the more complex the utterance).

3. *Accuracy* is "how well language is produced in relation to the rule system of the target language" (Skehan, 1996, p. 22). Specific grammatical features are investigated in previous studies, including regular and irregular past tense forms (Ellis, 1987), the use of definite and indefinite articles (Crookes, 1989), noun-modifiers and articles (Ortega, 1999), and plurals and verb morphology (Yuan & Ellis, 2003).

Generalized measures are also used by a number of studies (Skehan & Foster, 1997, 1999).

Limited attentional resources result in the prioritization of some aspects of language over others. Because of the competition of attentional resources among the three key aspects of speaking, tension builds up. This results in a trade-off effect for these three pedagogical goals, where enhancement in one aspect is due to lead to the detriment of another. Thus it is important to achieve an optimum balance among these goals. Previous research found that this trade-off effect occurs between complexity and accuracy (Foster & Skehan, 1996; Skehan & Foster, 1997, 1999; Tavakoli & Skehan, 2005), whereas other studies (Wendel, 1997, as cited in Yuan & Ellis, 2003; Yuan & Ellis, 2003) found that mutual tension exists between fluency and accuracy. This issue is discussed later in this section.

Concerning the issue of planning, Skehan's (1998) approach indicates that given the opportunity to plan before the task, learners are able to access their mental resources and thus the burden on learners' working memory is lessened, which in turn lessens the cognitive demand of the task. Thus planning has positive effects on learners' task performance and on the fluency and complexity of learners' language in particular.

The alternative approach to tasks is suggested by Robinson (2001). According to his view, the more cognitively demanding the task, the more accurate and complex the learner gets, whereas cognitively less demanding tasks promote fluency only. He

argues that learners are like native speakers, and have the capacity to attend to more than one aspect of language at the same time.

According to Robinson (2001), task complexity is determined by two sets of features. The first of these is the "resource-directing" factor (Robinson, 2001, p. 30), which includes the number of task elements, the reasoning demands within the task, and whether the immediacy of information is provided. The presence of "resource-directing" (Robinson, 2001, p. 30) factors in a task increases the cognitive demand on learners. In other words, these tasks offer no contextual support, have lots of reasoning demands and offer no pre-task planning time. In such cases, Robinson (1995, 2001) claims that learners will "rise to the challenge" (Robinson, 2001, p. 33) and draw on a greater range of syntactic resources than when they are given an easier task to do, thus enhancing both complexity and accuracy but not fluency.

Next is the "resource-depleting" factor (Robinson, 2001, p. 30), which includes of pre-task planning time (the opportunity to plan before task performance), the number of tasks the learner has to complete, and prior knowledge to the task. Tasks with a low cognitive demand involve "resource-depleting" factors, such as having contextual support, no reasoning demands and allowing for learners to plan for the task. This type of task promotes fluency at the expense of accuracy and complexity. By manipulating these factors, the cognitive demand required for task performance will vary, which in turn leads to the variation in quality of the language produced.

As evidenced in the discussion above, Robinson's (2001) and Skehan's (1998)

views towards pre-task planning conflict. If Robinson's (2001) model is followed, then pre-task planning should be viewed as a "resource-depleting" factor which promotes fluency but not accuracy or complexity. Non-planners who face "resource-directing" factors are forced to draw on a greater range of syntactic resources, thus producing narratives which are more complex than pre-task planners, whereas pre-task planners, on the other hand, face "resource-depleting" factors, which pose a detriment to their complexity level. The opportunity for pre-task planning becomes an obstacle to achieving more complex and accurate language. This contrasts with Skehan's (1998) claim that pre-task planning reduces the cognitive load of the working memory, allowing learners to allocate more attentional resources to target linguistic features, and thus is beneficial to complexity, if not always accuracy. It should also be noted that the majority of studies on pre-task planning have supported Skehan (1998)'s framework of task performance. The concept of pre-task planning is discussed in the next section.

After examining the underlying concepts to speech processing and different frameworks of tasks, the different types of planning in oral tasks, the different planning conditions under which the tasks are operationalized, and how they affect the fluency, accuracy and complexity of learners' language production, are discussed.

2.3 The Two Types of Task Planning -- Pre-task Planning and On-line Planning

Before moving on to discuss the two types of task planning, several points can be

concluded from the previous section. First, the information-processing theory shows that the working memory has limited capacity which only allows a limited amount of information to be processed at the same time. Thus bottlenecks are easily formed in the working memory when there is a lot of linguistic information to be processed, especially when the language learners are low in proficiency level, or when the task contains information which is unfamiliar to the learner.

Second, the limitations of capacity in the working memory, and the lack of attentional resources of language learners, shows that learners have to make decisions on allocating their attention to certain prioritized aspects of language. Without a focus-on-form in tasks, learners primarily prioritize message conveyance of a task and thus find it difficult to attend to both meaning and form at the same time.

Third, in Levelt's (1989) model of speech production, the two stages of formulation and articulation require controlled processing (as opposed to automatic processing) in L2 learners. This becomes a problem for low proficiency learners, who may find these processes taxing to their working memory in their retrieval of linguistic resources.

The attentional limitations that L2 learners face when dealing with an unfamiliar language task indicates the usefulness of task planning. Baddeley's (Baddeley & Hitch, 1974; Baddeley & Logie, 1999) model shows that rehearsal helps retain information in the brain, which lessens the burden on the working memory. Moreover, planning allows language learners to compensate for their lack of language expertise

and linguistic resources. In short, planning can help overcome the limitations of attentional resources in language learners. This is also supported by a number of studies focusing on task planning (Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Ortega, 1999, 2005; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003).

The positive effects of task planning on learners' language production fit with Skehan's (1998) cognitive approach to tasks, which states that planning before the task improves learners' language production, which is adopted for this study. According to Skehan (1998), tasks that do not give learners the opportunity to plan are labeled tasks with greater cognitive difficulty, while those providing learners with planning time are tasks with less cognitive difficulty. The tasks with less cognitive difficulty can lead to an improvement in the fluency and complexity of learners' language. In short, the importance of planning receives full support from the literature (Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Ortega, 1999, 2005; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003).

There are two major types of task planning. Ellis (1987) first draws our attention to different types of planning when he found that the accuracy of discourse differed in written narratives, re-done oral narratives and new oral narratives, after which followed a large amount of research on pre-task planning (Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Ortega, 1999, 2005;

Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003). Ellis (2005) divided task planning into 1) unpressured on-line planning and 2) strategic planning (pre-task planning).

According to Yuan and Ellis (2003), "on-line planning is the process by which speakers attend carefully to the formulation stage during speech planning and engage in pre-production and post-production monitoring of their speech acts" (p. 6). This definition relates directly to Levelt's (1989) model of speech processing. Allowing learners more time to complete a task is assumed to assist the planning and production of speech, by allowing the search for grammatical linguistic resources, facilitating the process of pre-production monitoring, and encouraging the process of post-production monitoring (Yuan & Ellis 2003).

The second type of planning, strategic planning, or pre-task planning, refers to providing learners with the opportunity to think about the task before performance. According to Skehan (1996), pre-task planning serves "to prepare the learners for handling the cognitive demands that a task contains and enable the channeling of attentional resources towards language form" (p. 25). Besides giving learners limited planning time, studies on pre-task planning (Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Ortega, 1999, 2005; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003) have manipulated planning conditions in order to test for optimum benefits for language production.

The number of studies which have explored pre-task planning is great when

compared with those that focused on on-line planning. This is due to the numerous ways to manipulate pre-task activities and planning conditions, while on-line planning is relatively difficult to operationalize. In order to be comparable with most of the studies on task planning, this study focuses on pre-task planning. The next section discusses the different pre-task planning conditions which are operationalized in previous studies and their influence on learners' language production.

2.4 Manipulating Task Characteristics under Different Planning Conditions

There is a vast amount of task characteristics that can be manipulated with pre-task planners, in order to vary the degree of task difficulty of the task. These include guided planning vs. unguided planning (Foster & Skehan, 1996), group planning vs. solitary planning (Skehan & Foster, 1999), personal tasks vs. narrative tasks vs. decision-making tasks (Skehan & Foster, 1997), different lengths of planning time (Mehnert, 1998), structured tasks vs. unstructured tasks (Tavakoli & Skehan, 2005) and language expertise (Ortega, 1999, 2005). How do these variables, together with pre-task planning, influence the key aspects (i.e. fluency, accuracy and complexity) of learners' language production?

Research has found that pre-task planning enhances the fluency of speech (Bygate, 1996; Crookes, 1989; Foster & Skehan, 1996; Ortega, 1999; Skehan & Foster, 1997). In Foster and Skehan (1996), half of the language learners were given instructions during pre-task planning as to which aspect of language they should focus

on. These learners were compared with those who had the opportunity to plan but were not provided with any instructions as to what linguistic aspects to focus on. Foster and Skehan concluded that learners performing the task under the guided condition produced more fluent speech. Mehnert (1998) compared learners who performed under four pre-task planning conditions: no planning, 1-minute planning, 5-minute planning and 10-minute planning. She concluded that fluency increases with the amount of planning time. However, the difference between learners' language production was greatest between non-planners and 1-minute planners.

Skehan and Foster (1999) attempted to find out whether tasks with a clear inherent structure, particularly in terms of time sequence, would lead to improved task performance. They used video prompts instead of the traditional picture prompts that task planning studies used, and concluded that planning helps learners access linguistic material from the memory more easily and rapidly.

Similarly, Tavakoli and Skehan (2005) made use of various picture prompts, and stated that these picture sequences contained varying degrees of structure in the storyline. They found that fluency increased with the opportunity to plan before the task *and* with the availability of a clear structure within the storyline. Ortega (1999, 2005) studied intermediate and advanced language learners of Spanish, and concluded that fluency is enhanced by pre-task planning.

Most studies (Bygate, 1996; Crookes, 1989; Foster & Skehan, 1996, 1999; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan

& Ellis, 2003) have also showed that pre-task planning has a positive effect on complexity. Crookes (1989) reported that learners use planning time to complexify the task, which results in a greater amount of subordination. Bygate (1996) explored the effects of task familiarity with video prompts, where learners were required to re-tell the same story from the video with a gap of three weeks in-between. He concluded that with the increase of familiarity of the task, complexity is increased with a more complex argument structure in verb phrase, nominalization and a clearer quality of lexical selection.

Mehnert (1998) found that greater complexity is evident for the 10-minute planner group in her study, as compared with non-planners, 1-minute planners and 5-minute planners. Foster and Skehan (1999) examined the difference between learner planning in groups and planning individually. They found that individual learner planning produces the best effects for complexity.

Ortega (1999), who studied advanced learners, reported that planners produce longer utterances. Yuan and Ellis (2003) explored different types of planning – pre-task planning and on-line planning – and found that pre-task planning had strong positive effects on syntactic complexity. Foster and Skehan (1996), Wendel (1997), Wigglesworth (1997), Skehan and Foster (2001) and Tavakoli and Skehan (2005) also came to the same conclusion.

Unlike fluency and complexity, the research (Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Ortega, 1999, 2005; Skehan &

Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003) has generated mixed findings for the effects of pre-task planning on accuracy and is therefore inconclusive. Ellis (1987) required learners to produce written narratives, then re-do the same story in oral narratives, and also produce new oral narratives all in a row. He reported that planning had positive effects on the accuracy only on regular past tense forms, but not on the irregular past.

Foster and Skehan (1996) studied different types of tasks, including personal tasks which are based on information familiar to the learners, narrative tasks in which learners have to tell a story based on picture prompts, and decision-making tasks which "were more interactive and which required the capacity to relate a set of moral values to a series of decisions that had to be made" (p. 98). These tasks were operationalized under guided and unguided planning conditions. Guided planners were given detailed instructions as to which target linguistic forms to plan before the task, whereas these instructions were not provided to unguided planners. Foster and Skehan (1996) found that while pre-task planning did have an impact on accuracy, the condition had to be unguided planning.

Skehan and Foster (1997), who also explored these three task types, found that accuracy was enhanced with pre-task planning when combined with personal or narrative tasks, but not with decision-making tasks. Mehnert (1998) found that the accuracy of 1-minute planners benefits from pre-task planning, but that this was not the case for 5-minute or 10-minute planners.

Ortega (1999) reported that accuracy is enhanced with pre-task planning on noun-modifiers, and a follow-up study reports that pre-task planning has a positive effect on articles (Ortega 2005). Skehan and Foster (1999) and Tavakoli and Skehan (2005) found accuracy improved with the combined effects of pre-task planning and structure, but that proficiency level had the greatest effect. Iwashita et al. (2001) even reported that accuracy increases under more cognitively demanding conditions. In contrast to this, Crookes (1989), Wendel (1997), Wigglesworth (1997), Skehan and Foster (2001) and Yuan and Ellis (2003) all showed that pre-task planning has no positive effects on accuracy.

In other words, accuracy is influenced by a number of factors, including the type of planning (accuracy is increased with unguided planning), the type of task (personal or narrative tasks favour accuracy), the grammatical feature explored (when the feature involves a clear rule, accuracy is enhanced), task complexity (a task which is cognitively less demanding favours accuracy), task structure (structured tasks favours accuracy), the length of planning time (limited planning time leads to increased accuracy), and learner proficiency (advanced learners are able to benefit more from planning).

However, it is obvious from the studies mentioned above that the findings on accuracy are not only inconsistent among different studies, but that within the same study accuracy is also able to produce mixed results. The above generalizations, despite being drawn from previous studies, are by no means as simple as they seem.

Their interaction with each other and how this influences accuracy, are issues that remain to be examined.

In addition to this, learners with limited processing capacity have to prioritize their pedagogical goals when speaking, thus there is an additional factor to accuracy effects in learners' language production -- the trade-off effect among fluency, accuracy and complexity. Studies (Crookes, 1989; Foster & Skehan, 1996, 1999; Mehnert, 1998; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003). show that the trade-off effect possibly involves complexity and accuracy. Results from Crookes (1989) showed that while there is a strong effect of pre-task planning on fluency and complexity, accuracy remains unchanged. Foster and Skehan (1996) reported that guided planners complexify the task, thus they are less accurate. Skehan and Foster (1997) concluded that planners prioritize either complexity or accuracy, but not both. Mehnert (1998) showed that 1-minute planners devote their attention to accuracy instead of complexity within the limited time, while 10-minute planners do the opposite.

Tavakoli and Skehan (2005) found that accuracy and complexity are both aspects of form, which means that strong mutual tension exists between them. Where learners were willing to take risks with using more complex syntax or lexis, accuracy was compromised, whereas when learners tended to conform to grammatical rules, syntactic complexity and lexical variety were neglected. Thus for the studies mentioned above, the trade-off effect involves accuracy and complexity -- where

complexity is enhanced, accuracy is sacrificed.

However, there are also studies showing that the trade-off involves accuracy and fluency. Wendel (1997, as cited in Yuan and Ellis, 2003) suggested that the type of planning determines whether learners predispose their attention to accuracy or fluency, and that with pre-task planning, learners attend to fluency, whereas the increase in accuracy is involved with online planning (but it must be noted that he made no comment on complexity.) Results from Skehan and Foster (1999) also showed that while there is an increase in fluency when there was a predictable structure in the storyline, there is no effect on accuracy or complexity, suggesting that fluency and accuracy might be in tension. It can be concluded that multiple factors affect pre-task planners. Yuan and Ellis (2003) also displayed results which confirmed Wendel's (1997) claim. Though firm conclusions can be drawn on fluency and complexity, accuracy still remains to be a debatable issue.

Following the literature mentioned above, several generalizations of pre-task planning studies can be drawn. First, narrative tasks are generally used to investigate pre-task planning effects (Foster & Skehan, 1996, 1997; Ortega, 1999; Tavakoli and Skehan, 2005; Skehan & Foster, 1997; Wigglesworth, 1997; Yuan & Ellis, 2003). Second, these narrative tasks are usually based on picture prompts extracted from writing books (Foster & Skehan, 1996, 1997; Ortega, 1999; Skehan & Foster, 1997; Wigglesworth, 1997).

Only a limited number of studies test learners with video prompts (Bygate, 1996;

Skehan & Foster, 1999). Third, most of these studies focused on investigating quantitatively, with little room for an in-depth exploration of learners' perspective. Ortega (1999, 2005) is the only exception. With reference to these issues, this study attempts to devote more attention to some of the areas in which previous literature on pre-task planning has neglected. These aspects include the use of alternative narrative prompts and the focus on learners' perception of task difficulty.

2.5 The Two Types of Narrative Prompts – Picture Sequences and Video Snippets

The benefit of using video prompts lies in the clear inherent structure of the clip. Because of this, task structure will be explained in greater detail in this section. Levelt's (1989) model of speech production (see previous section) shows that an available script in the conceptualizer reduces its pressure and results in more attentional resources allocated to the formulator, which in turn benefits fluency and accuracy. Kobayashi and Hoey (1993) identified the relevance of a problem-solution structure in a narrative task, which helps learners improve task performance. Foster and Skehan (1996) and Skehan and Foster (1997) stated that components of a macrostructure include a clear timeline, a script, and a story with a conventional beginning, middle and end. In addition, it has to appeal to what is familiar and organized to the learner.

A number of studies using structured picture prompts found that these tasks benefit fluency and accuracy (Foster & Skehan, 1996; Skehan & Foster, 1997, 1999;

Tavakoli & Skehan, 2005). Foster and Skehan (1996) and Skehan and Foster (1997) explored 3 task types (narrative, personal and decision-making tasks) and found that tasks with a clear macrostructure benefit fluency and accuracy but not complexity. Skehan and Foster (1999), who utilized video snippets from *Mr. Bean* as the main testing material, found that when there is a predictable structure in the task, fluency and accuracy benefit while complexity remains unaffected. Tavakoli and Skehan (2005), who hypothesized that task structure should not be a dichotomy but a continuum, confirmed the positive effect of task structure on fluency and accuracy, except that they also found that complexity scores relatively high on the second most-structured task. From this we can generalize that an inherent structure in narratives favours fluency and accuracy.

The studies mentioned above draw our attention to the traditional use of prompts in pre-task planning research, which shows that picture sequences are the major testing material in narrative tasks. Ellis (1987) utilized picture prompts to find out how planned discourse influenced accuracy, especially focusing on the simple past tense in English. Skehan and Foster (1997) used a cartoon strip from Sempe, a French cartoonist. Ortega (1999) extracted pictures from Hill (1960). Yuan and Ellis (2003) made use of a pictures series from Heaton (1975). Other studies that made use of picture sequences include Wigglesworth (1997), Iwashita et al. (2001) and Tavakoli and Skehan (2005). Obviously pictures sequences do provide a context for participants to tell a story, and when used to elicit language performance, are justified

in terms of construct validity, reliability, and the authenticity of tests.

A limited number of studies made use of videos as the prompt for narrative tasks. This includes several studies that explore task repetition (Bygate, 1996), task structure (Skehan & Foster 1999), and carryover effects (Gass & Mackey 1999). Bygate (1996) used *Tom and Jerry* video segments of a minute and a half in length to measure linguistic complexity. Skehan and Foster (1999) made use of *Mr. Bean* video segments as the testing material for both structured and unstructured tasks. Gass and Mackey (1999) extracted *Mr. Bean* video snippets to see whether more accurate language or sophisticated language use carried over to a new context.

One other study which utilized picture prompts for narrative tasks found that their insignificant results were due to the narrowness of picture prompts. Iwashita et al. (2001), a study which has failed to show significance in the manipulation of various task characteristics (perspective, immediacy, adequacy and planning time), suggested a remedy for their failure to elicit significant results in the aspect of complexity. They suggested that involving learners in "recasting more complex output of the kind, for example, provided by television with 'soaps' containing dialogic sequences" (p. 429).

Despite the overwhelming use of picture prompts in narrative tasks, Bygate (1996) found that the use of videos increased learners' provision of "scene-setting information," not to say providing the motivations and intentions for video characters, and their "abstracts" of what is about to happen. Although Skehan and Foster (1999) and Gass and Mackey (1999) did not specify any benefits derived from the difference

between using pictures and using videos as prompts, the appropriateness of video snippets as representations of structured tasks cannot be denied.

Moreover, studies to date on pre-task planning have never compared the use of picture prompts and video prompts in narrative tasks, and whether there is a difference in the elicited language performance. Do videos, which are so different from pictures with their audio-visual effects and possibly much more dialogue, elicit similar levels of fluency, accuracy and complexity as pictures? As videos also provide much more information (and less room for imagination) on the storyline than pictures, does this make a difference to learners' language production? Do learners find videos more enjoyable and less difficult to interpret? Does a change in the learner's attitude affect their task performance?

2.6 Learners' Perception of Task Performance

Over the past decade, research (Bygate, 1996; Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Iwashita et al., 2001; Ortega, 1999, 2005; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003) on pre-task planning has focused on manipulating task characteristics from the *researcher's* perspective. Tasks are designed according to what the researcher thinks is appropriate, but seldom is the perspective of the *learner* taken into account. Ortega (1999, 2005), Elder et al. (2002), Yuan and Ellis (2003) and Tavakoli and Skehan (2005) are some of the studies that include the qualitative dimension into their study, and thus shed light on learners'

perception of the usefulness of planning, the difficulty of the task and the relationship between learner attitudes and task performance.

Post-task questionnaires and retrospective semi-structured interviews are the two major methods used in these studies. Ortega (1999, 2005) aimed to find out what learners do when they plan, and how they allocate their limited attentional resources. Through post-task interviews she found that language expertise is a factor for differences in task performance, although it does not lead to much difference in strategy use during pre-task planning. Both intermediate and advanced learners reported using retrieval and rehearsal operations; however, advanced learners were able to balance these two operations, while intermediate learners were not. Self-monitoring strategies were also reported by advanced learners, whereas their intermediate counterparts were more concerned with the retrieval of grammatical items instead. These two differences helped explain the higher syntactic and lexical complexity, and also accuracy, of advanced learners over intermediate learners.

In addition to this, it is important to point out that SLA research in task planning so far has provided very little information about what learners do while they plan. To date, only a limited number of studies, including Ortega (1999, 2005), investigated what learners do when they are given the opportunity to plan. This is a potentially important factor that must be taken into account of before coming to any conclusion on task performance. It is still unclear whether learners think that planning is useful and what the strategic processes are, that underlie pre-task planning. More

interestingly, is there a positive relationship between learners' perception of task difficulty, or learner attitude, and task performance?

Ortega (1999) also found that the proficiency level of learners affected what was planned. Planning was more likely to have an effect on the quality of linguistic output of advanced learners. Similar to Wigglesworth (1997), she thought that the proficiency level of learners was a moderating factor that needed to be included in previous research, especially concerning studies that find the lack of accuracy effects (Wendel, 1997) and the trade-off hypothesis between accuracy and complexity (Foster & Skehan, 1996; Skehan & Foster, 1997). Tavakoli and Skehan (2005) also found that fluency and accuracy had a significant difference between advanced and intermediate learners, whereas there was a significant interaction between the task and proficiency level for complexity.

Tavakoli and Skehan (2005) also investigated language expertise using post-task questionnaires. Yuan and Ellis (2003) also utilized the same method to find out how learners allocate their attentional resources during pre-task planning. Neither of these studies, however, attempted to link learners' perception to actual task performance.

The only study that explored this issue is Elder et al. (2002). Their study focused on the two items of task difficulty and task enjoyment. They found that learners' perception of task difficulty only revealed significant relations with some of the task dimensions (perspective and adequacy and planning in a certain task, but not immediacy). As for task enjoyment, although learners showed higher enjoyment under

certain conditions (immediacy), the effect size and the difference with performance under other conditions were small. Because all of them had weak correlations, Elder et al. (2002) came to the conclusion that the relation between task performance and test-taker attitudes cannot be sustained.

There have been claims that learners' attitudes are significantly related to their task performance (Bradshaw 1990; Brown 1993; Shohamy 1982; Zeidner 1988, 1990). Brooks (1999, as cited in Elder et al., 2002) also reported a relationship between attitude towards different assessment types and levels of performance on respective tasks. Robinson (2001) compared the reactions of ESL learners to 2 different versions of the same task. He found that not only was complexity significantly influential on various aspects of learner production, but it was also associated with learner perceptions of task difficulty.

Drawing from the studies above, it is found that performance differences related to language expertise can be elicited through qualitative rather than quantitative means. It can also be generalized that learner proficiency is indeed an important mediating factor to task performance. This leads to the questions of: How does proficiency level affect the task performance of planners? Do video narratives elicit different results from advanced and intermediate learners, as in research that has investigated narrative tasks on traditional picture prompts?

2.7 Research Questions of This Study

Drawing from the literature above, the following research questions are asked.

1. What effect does pre-task planning, in addition to the use of picture prompts and video prompts, have on the *fluency* of L2 learners' production in an oral narrative task?

In accordance with the preceding discussion, pre-task planning is expected to result in greater fluency of language use.

2. What effect does pre-task planning, in addition to the use of picture prompts and video prompts, have on the *complexity* of L2 learners' production in an oral narrative task?

It is predicted that pre-task planning will elicit language with greater complexity.

3. What effect does pre-task planning, in addition to the use of picture prompts and video prompts, have on the *accuracy* of L2 learners' production in an oral narrative task?

Previous studies have reported mixed results for the effects of pre-task planning on accuracy. It is predicted that no effect will be found in this study because the pre-task planning group have limited opportunity to plan online.

4. What are learners' perceptions of the usefulness of pre-task planning time, task enjoyment or task difficulty? What do learners do when they plan?

It is predicted that learners, especially intermediate learners, will find pre-task planning time useful for task performance. Learners will also find that tasks which offer the opportunity to plan beforehand less difficult than those without the

opportunity to plan. As for task enjoyment, it is predicted that learners will find greater task enjoyment with tasks on video prompts, rather than picture prompts. Concerning what learners do when they plan, Ortega (1999, 2005) has reported different planning strategies among advanced and intermediate learners. It is predicted that learners will generally focus on retrieval and rehearsal operations during planning.

The next chapter describes how this study operationalizes pre-task planning time with both picture and video prompts, and the post-task methods used to elicit learner perspective of task difficulty.

CHAPTER THREE

METHOD

3.1 Introduction

This study is a 2 x 2 design which manipulates two between-participant variables with two levels of conditions each. These two variables include: narrative prompts (picture sequences and video snippets) and planning time (no planning and pre-task planning). 16 participants were given a pre-task questionnaire about their language background to ensure that their English proficiency matched the purposes of the test. After that, they were equally divided into the two proficiency groups, where each group performed 2 oral narrative tasks elicited by means of the prompts mentioned above, under designated planning conditions.

3.2 Participants

The participants in this study were 16 full-time students who majored in English at the Chinese University of Hong Kong. Their age range was 20-23, all of them being Hong Kong Chinese whose mother tongue was Cantonese.

At the time when the data of the present study were collected, most of these learners had been learning English as a foreign language for at least 17 years, first at kindergarten, then at primary school, secondary school and then at the university. Most of them studied in Chinese schools which were labeled English Medium of

Instruction (EMI) schools, meaning that almost all lessons were taught in English. A few of the participants studied in Chinese Medium of Instruction (CMI) schools, where almost all lessons were conducted in Cantonese. About half of them had traveled abroad to an English-speaking country, but other than that, few of them had the opportunity to use English for communicative purposes outside the classroom. Despite this, they were regularly exposed to English through the media, such as television, movies, and pop songs, hence more than half expressed comfort when conversing in English. Refer to Table 1 below.

As undergraduates, all participants had taken the Advanced-Level Use of English Examination (HKALE: UE) before admission into the university. Their general scores ranged from A to D in this exam, with their oral exam also taken into consideration. Hence two groups of participants were formed: 6 females and 2 males with grades A or B were allocated into the advanced learners group, while the rest consisting of 5 females and 3 males with grades C or D were allocated into the intermediate learners group. At the university, all of them took at least 4 English courses per semester (i.e. 12 hours of English courses per week). Each 2-hour lecture comes with a tutorial, during which students' active participation in class is required.

Table 1

Participant information

Participant#	Gender	Age	Overall grades in UE	EMI / CMI*	# of Eng courses	Exposure outside classroom	Traveled to English-speaking countries?	Comfortable with using Eng?
1	F	20	B	EMI	4	books, movies, pop music	Yes	average
2	M	21	D	CMI	4 <	internet, pop music	Yes	not very
3	F	20	B	EMI	4 <	books, internet, movies, pop music	No	very
4	M	22	C	CMI	4 <	pop music	Yes	not very
5	F	23	A	EMI	3	books, chat with foreigners, online blog	Yes	average
6	F	21	D	EMI	4 <	books, music, movies, chat with foreigners	No	very
7	F	22	A	EMI	4 <	books, movies	No	average
8	F	20	D	EMI	4	emails, tv, penpals,	Yes	very
9	F	21	B	EMI	4 <	books, diary, homework, lectures	Yes	extremely
10	F	20	D	EMI	4 <	part-time job	No	not very
11	M	21	B	EMI	4 <	blog, chat with foreigners, music	Yes	extremely
12	F	23	C	EMI	4	internet, magazine, tv	Yes	average
13	F	23	B	EMI	4	assignments, chat with foreigners	Yes	extremely
14	M	22	C	CMI	3	music, newspapers, tv	Yes	average
15	M	22	B	EMI	4	tv, online blog	Yes	extremely
16	F	21	C	EMI	4	books, movies, newspapers	No	not very

*EMI / CMI = Schools using English / Cantonese as the medium of instruction for teaching respectively

Undergraduates in their second and third years were invited and to participate in the study. They were told that the test and tasks they would complete were for the purposes of research but were not told the precise purpose. They were also assured that the information collected would not be used towards their course grades.

3.3 Task Type

The task required participants to narrate a story orally based on a picture composition from Huizenga (2000) and a video snippet from *The Simpsons* (2000). These types of tasks were chosen because, first, similar tasks have been used in other studies of planning (Bygate, 1996; Foster & Skehan, 1996; Iwashita et al., 2001; Skehan & Foster, 1997, 1999) and thus comparison with the results of these studies would be easier. Individual monologic narrative tasks also involve only a single type of stimulus of the kind routinely used by international testing organizations such as Test of English as a Second Language (TSE), thus they are justified in terms of the reliability and authenticity of the test. Second, oral narratives are monologic rather than dialogic, thus affording a basis for deriving measures of learner performance that are not influenced by interactional variables.

This study used alternative prompts for narrative tasks, in addition to the traditionally-used picture sequences. Previous studies on planning (Foster & Skehan, 1996; Iwashita et al., 2001; Skehan & Foster, 1997) have focused on picture sequences as the major testing tool for narrative tasks as this allows the task to be

reasonably demanding by requiring interpretation on the part of the learners (Skehan & Foster, 1999). Only a limited number of studies used alternative prompts, such as video snippets (Bygate, 1996; Skehan & Foster, 1997). One of these studies, Bygate (1996), found that with video prompts, participants increased their syntactic complexity and metalinguistic comments after doing tasks. Because of the lack of research on narrative prompts other than picture sequences, and to find out the comparability of learners' language elicited by picture sequences or video snippets, both types of narrative prompts were used in this study.

4.4 Pilot Studies

3.4.1 Pilot Study on Picture Prompts

Two pilot studies were carried out to test the appropriateness of prompts. The first pilot study tested the appropriateness of picture sequences extracted from Heaton (1975), one of the most popular source used by pre-task planning studies (Ellis, 1987; Foster & Skehan, 1997, 1999; Wigglesworth, 1997; Yuan & Ellis, 2003). The alternative source of picture sequences is a writing textbook published in the last few years: Huizenga (2000). The second pilot study tested for video snippets and compared participant reaction towards clips from *Mr. Bean*, which had been previously used in Gass and Mackey (1999) and Skehan and Foster (1999), and a clip from *The Simpsons*.

The first pilot study compared participant reaction towards the two selected

picture sequences. For studies on planning that use narrative tasks as the tool, Hill (1960) and Heaton (1975) are the two most commonly used sources for picture sequences in task-based studies (Ellis, 1987; Foster & Skehan, 1997, 1999; Wigglesworth, 1997; Yuan & Ellis, 2003). However, these two sources are designed for second language learners at the beginners' level (Heaton is even used as guided writing material in some primary schools in Hong Kong). Moreover, both are clearly outdated in the trend of education textbooks. Thus it is questionable whether the two could serve as the main testing material for participants obtaining a university degree in the target language. To correct the possible misrepresentation of their skills and test for the appropriateness of this material, a pilot study was carried out. To ensure comparability with the literature (Iwashita et al., 1998; Ortega, 1996, 1999; Tavakoli & Skehan, 2005), the picture sequence has to contain 1) 6-8 pictures, 2) a complete storyline, and 3) an inherent structure.

This pilot study was carried out among 4 English majors with a language background comparable to that of the main participants. They did narrative tasks with both Heaton (1975) and Huizenga (2000) (see Appendix A). All four participants displayed positive interest in Huizenga (2000) as it involved more depth in the topic and a more complicated structure, whereas the picture sequence from Heaton (1975) elicited only simple grammar and vocabulary because of its straightforward storyline. All participants agreed that Huizenga (2000) is more appropriate to be used in the main study.

3.4.2 Pilot Study on Video Prompts

A second pilot study was carried out to test for the appropriateness of the video snippet. For previous studies using video prompts, *Tom and Jerry* cartoons (Bygate, 1996) and *Mr. Bean* movies (Gass & Mackey, 1999; Skehan & Foster, 1999) have been used. The criteria for choosing video clips are: 1) there must not be too much dialogue; 2) the stories must be condensed into limited time frames with a complete situation-problem-evaluation structure (Kobayashi & Hoey, 1993), and 3) preferably the story should focus on a single character.

Mr. Bean movies have been chosen by previous studies (Gass & Mackey, 1999; Skehan & Foster, 1999) not only because it satisfies the criteria listed above, but also because the program is also available on television worldwide, which is assumed to ensure an easy understanding and free interpretation on the part of the participant.

Despite this, a previous study (Mok, 2005, unpublished paper) found that *Mr. Bean* does not interest participants from Hong Kong at all. From this study, which looked for the effect of different planning conditions on Hong Kong university students, I concluded from the participants' response that the above criteria gave great constraints to their enjoyment of the video, which, according to Elder et al. (2001), which explored test-taker attitudes, might affect task performance. Following this, *The Simpsons* was chosen for comparability with a *Mr. Bean* video clip. The same participants in the first pilot study were invited to participate in this second pilot

study.

For the second pilot study, all of the participants agreed that the episode from *The Simpsons* was more enjoyable, both while watching the clip and retelling the story. Not only did it arouse and capture the enthusiasm of the participants, but participants also expressed comfort with it despite it being an American cartoon, whereas there was a general failure to understand the British humour in *Mr. Bean*. All of them agreed that *The Simpsons* was the more appropriate choice for the main study. The selected picture sequence from Huizenga (2000) and video snippet from *The Simpsons* are described in the next section.

3.5 Task Design

Following the 2 pilot studies, a picture story from Huizenga (2000) and a video snippet from *The Simpsons* (2000) were chosen as the main testing material for the study. The picture sequence from Huizenga (2000) is a story about a lonely man and woman living in different countries who met on the internet, fell in love with each other, and despite their families' objection, became a couple in the end (see Appendix A). The video snippet from *The Simpsons* (2000) is titled *Hell Troupe*. The story opens with a criminal called Snake getting justice and receiving the death penalty, but unluckily Homer Simpson has inherited Snake's vengefulness with the transplantation of Snake's hair onto his head. He went on to murder two witnesses who had sent Snake to the electrical chair, and was about to strike his son, Bart, when his moral

strength overcame the evil of the hair. Both stories can be considered demanding given that it was necessary for the participants to distinguish between the different characters in the two stories, and also to interpret the message of the story.

As mentioned in the previous section, the 16 participants were divided into 2 groups, advanced and intermediate, according to their English proficiency. Under the two planning conditions and the two prompt types, 4 combinations of tasks were formed. Please refer to Table 2 below. The four tasks were: 1) no planning with pictures, 2) planning with pictures, 3) no planning with videos, and 4) planning with videos.

Table 2

Tasks Manipulated under Four conditions

Prompt type	**Pre-task planning**	
	0.5 min (no planning)	3.5 min (limited planning)
Pictures	P1	P2
Videos	V1	V2

Each participant completed two tasks in a row. 4 groups were further divided

from the two proficiency groups so that each participant had the opportunity to complete a task on both pictures and videos. Planning conditions were also randomized so that the 4 tasks were randomized into 4 sets, as shown in Table 3.

Table 3

Different Sets of Tasks

	Set(s)			
	1 (n=4)	**2 (n=4)**	**3 (n=4)**	**4 (n=4)**
Task 1	P2	P1	P2	P1
Task 2	V2	V2	V1	V1

Each set of four participants was made up of two advanced learners and two intermediate learners. According to Table 3, four participants from Set 1 performed the two tasks under the conditions *planning, pictures* and *planning, videos*. Participants belonging to Set 2 performed the two tasks under the conditions no *planning, pictures* and *planning, videos*. Participants belonging to Set 3 performed the two tasks under the conditions *planning, pictures* and *no planning, videos*. Participants belonging to Set 4 performed the two tasks under the conditions *no planning, pictures* and *no planning, videos*. The order of the tasks were counterbalanced (i.e. half of the participants did the task with pictures first, then

videos). Refer to Table 4 below.

Table 4

The Order of Task Performance

	Set(s)							
	1 (n=4)		2 (n=4)		3 (n=4)		4 (n=4)	
	1A (n=2)	1B (n=2)	2A (n=2)	2B (n=2)	3A (n=2)	3B (n=2)	4A (n=2)	4B (n=2)
1st Task	P2	V2	P1	V2	P2	V1	P1	V1
2nd Task	V2	P2	V2	P1	V1	P2	V1	P1

Specifically, the four participants from each set were split into half. Each subset, A and B, contained one advanced learner and one intermediate learner respectively. Participants from Subset A performed the tasks in the order of *pictures,* then *videos*. Participants from Subset B performed the tasks in the order of *videos,* then *pictures*. This was done so that the practice effect was minimized.

These tasks were designed to reflect different levels of cognitive difficulty. Please refer to Table 5 below.

Table 5

Task Design with assumed cognitive difficulty of each task

Task code	**Prompt**	**Planning (in minutes)**	**Requirements of dimension**	**Predicted difficulty (according to assumed degree of cognitive demand)**
P1	Pictures	3.5	tell a story from the pictures, with 3.5 minutes to plan beforehand	less difficult (lighter cognitive demand)
P2	Pictures	0.5	tell a story from the pictures, with 0.5 minutes only for looking at them	more difficult (heavier cognitive demand)
V1	Video	3.5	tell a story after watching the videos, with 3.5 minutes to plan beforehand	less difficult (lighter cognitive demand)
V2	Video	0.5	tell a story after watching the video, with 0.5 minutes only to recollect any thoughts	more difficult (heavier cognitive demand)

The time intervals were chosen to ensure comparability with the task-based literature that had an assessment focus. Under the no planning condition, tasks are assumed to have a heavier cognitive demand on participants, whereas tasks which allow for participants to plan have a lighter cognitive demand. As for prompts, picture sequences and video prompts are not supposed to make a difference to the difficulty of the task. However, video prompts are assumed to elicit more complex language from the participants than pictures sequences.

3.6 Data Collection Procedures

Data was collected under testing conditions on a one-to-one basis. Before the test, participants were familiarized with the objective and procedure of the test briefly in their mother tongue, i.e. Cantonese. Because of the importance of the listener role, which would make the task genuinely communicative (Ortega, 1999), participants were required to tell the story in the third person perspective, as they would to a friend who has not seen the picture sequence / watched the video clip beforehand. They also had to begin the story by "This is a story about….", following the example of previous studies (Ortega, 1999; Yuan & Ellis, 2003).

Participants were also told to take notes as they wished, but that note-taking was optional and the notes would not be collected. They were also not allowed to refer to either the notes, the video, or the pictures while re-telling the story. They were also

told to speak for approximately 3 minutes for each task, and notified that they would be recorded. Other than the above, no other instructions as to how to complete the tasks, or the detailed focus of the test, were given (see Appendix B).

The whole data collection procedure for each participant took about 45 minutes. Before the test started, all participants completed a language background questionnaire (see below). No placement test was used. After that, all participants had to complete 2 tasks. After completing and recording each task, participants completed a task difficulty questionnaire. Field notes were taken while the participants were producing speech data.

3.7 Pre-task / Post-task Questionnaires

A language background questionnaire which aimed to control the participants' proficiency level and their exposure level to English, was administered before the test (see Appendix C). Before the participants completed either questionnaire, each item was explained by the researcher in the participants' mother tongue, which is Cantonese. Some of the open-ended questions included:

Item 5. What courses did you take this semester that involves participation in English?

Item 6. How much are you exposed to English outside classroom? Please list occasions for reading, writing, listening or speaking.

Participants were also asked to rate their level of comfort when using English both inside the classroom (formal setting) and outside the classroom (informal setting). This was done using a 5-point Likert scale, where 1 represents the least comfortable, and 5 represents the most comfortable with using English. The results of this questionnaire are summarized in Table 1 (Participant Information), given above.

After each task, retrospective questionnaires were given to the participants for the triangulation of data (see Appendix D). These questionnaires aimed to find out participants' perception of task difficulty. Questionnaire items followed the examples from the literature (Ortega, 1995, 1999; Elder et al., 2002; Yuan & Ellis, 2003) (see Appendices C and D). There were specific question items on whether participants think that having pre-task planning time is useful for improving task performance, and whether they thought that the time provided for planning was adequate. For these question items, participants have to circle the appropriate response (yes or no) (see Appendix D).

To obtain an in-depth understanding of the participants' feelings, open questions were included. On the whole, question items were designed to tap only their reactions to the story-telling experience, without drawing their attention to the particular condition under which the tasks were performed. Some examples of open-ended questions are as follows:

Item 1. How did you spend your planning time? Did you think about grammar? Vocabulary? The best way to organize your story? Give examples.

Item 6. What would help you achieve better performance? You can target at the difficulties that you encountered when telling the story. Did you experience any difficulty in grammar? Vocabulary? Or how to organize your story?

Participants were also asked if they enjoyed the tasks, and whether they found the tasks easy or difficult. A 5-point Likert scale was used where participants had the choice to rate their response. For task difficulty (item 4), 1 represents that the task is least difficult whereas 5 represents that the task is most difficult. For task enjoyment (item 5), 1 represents that the task is least enjoyable whereas 5 represents that the task is most enjoyable (see Appendix D).

3.8 Quantitative Measures

Measures of accuracy, fluency and complexity were developed to evaluate the quality of the participants' oral production. These measures were largely the same as those used in other studies (Crookes, 1989; Foster & Skehan, 1996; Iwashita et al., 2001; Skehan & Foster, 1997, 1999).

3.8.1 Fluency measures

Fluency was measured according to speed, breakdown disfluencies (unfilled pauses, filled pauses), and repair disfluencies (false starts, repetition, reformulation, and replacement).

1. Speed: the number of words in each narrative task divided by the length of narrative (in seconds), times 60 (in seconds).
2. Unfilled pauses: the length of silent pauses at both mid-clause and clause ends. According to Freed (2000), these are disfluent-sounding silences which tend to be of 0.4 second or longer in duration.
3. Filled pauses: the number of meaningless words (e.g., "um", "er") both at mid-clause and clause ends.

As for the various sub-dimensions of repair disfluencies, the definition from Skehan and Foster (1999) is followed, as outlined below:

4. False starts: the number of utterances that are abandoned before completion.
5. Repetition: the number of immediate and verbatim repetitions of a word or phrase.
6. Reformulation: the number of phrases or clauses that are repeated with some modification either to syntax, morphology, or word order.
7. Replacement: the number of lexical items that are substituted for another. This can be either within the same clauses, within a following cause if this is otherwise a verbatim repetition, or within a following clause if this is a reformulation (Skehan & Foster, 1999).

3.8.2 Complexity measure

Syntactic complexity: the ratio of the total number of clauses to the total number of Analysis of Speech Units (A-S units). The A-S unit was introduced by Foster et al.

(2000) to compensate for the drawback of measures like T-units and C-units. Under their definition, an independent clause will be "minimally a clause including a finite verb, while a subordinate clause will consist of a finite or non-finite verb element plus at least one other clauses element (Subject, Object, Complement, or Adverbial)" (Foster et al., 2000, p. 365). Thus the transcribed data was coded into A-S units containing independent clauses and subordinate clauses.

3.8.3 Accuracy measure

Error-free clauses: the percentage of clauses that did not contain any grammatical or morphological errors. General measures of accuracy are used, which, though being blunt instruments, have been shown by Mehnert (1998) to generate significance, and do capture more variance in performance (Tavakoli & Skehan, 2005). All error-free clauses were identified and coded in the transcribed data, and the ratio of error-free clauses to the total number of clauses was calculated.

3.8.4 Data Analysis

The data was coded, with 10% of the data being coded by an independent expert against which the data coded by the author was tested. The inter-rater reliability coefficients were all above 0.90 for all codings. After that, a series of independent t-tests and two-way ANOVAs were performed on all measures.

3.9 Qualitative Measures

Questionnaire items focusing on learner perception of pre-task planning were categorized following Ortega (1999) and divided under two main groups: *Benefits of pre-task planning identified by learners* and *Limitations of pre-task planning identified by learners*.

Questionnaire items which asked about what learners plan, the difficulties they face when performing the task, and their perception of what would help them perform better, were coded as new categories emerge from the participants' response.

Last of all, learners' perception of task difficulty and task enjoyment were rated according to a 5-point Likert scale, following Elder et al. (2002). These items were subsequently analysed with a series of independent t-tests and two-way ANOVAs.

CHAPTER FOUR

RESULTS

4.1 Introduction

This section first explores the descriptive statistics obtained from learners' language production. The means and standard deviations of the various measures of fluency are given, including speed, pauses (filled), pauses (silence), false starts, reformulation, repetition, and replacements. The means and standard deviations of complexity and accuracy are also presented. These independent variables were also intercorrelated to look for significant interactions, followed by a series of independent t-tests and two-way ANOVAs.

A series of independent t-tests and two-way ANOVAs were also run for learners' perception of task difficulty and task enjoyment. Finally, learners' response to open-ended items in the retrospective task difficulty questionnaires was examined in detail. These items included whether learners thought that the availability of pre-task planning time was useful or adequate, how learners spent their planning time, and learners' opinion of what would help them improve their task performance.

4.2 Descriptive Statistics

This section first presents the descriptive data on the levels of performance achieved on the two tasks, and over the four conditions: 1) *planning, pictures*, 2) *no*

planning, pictures, 3) *planning, videos*, and 4) *no planning, videos*. This is followed by the main analysis used – a two-way analysis of variance (ANOVA) – which looked for significant interactions between the independent variables, pre-task planning and prompt, with the three key aspects of speaking: fluency, accuracy and complexity.

4.2.1 Overall Means and Standard Deviations

Table 6 shows the means and standard deviations of the measures of fluency, complexity and accuracy.

Table 6

Basic Descriptive Statistics on the Measures Used

Variable	*M*	*SD*
Speed	127.45	34.52
Pauses (filled)	13.19	16.11
Pauses (silence)	16.91	7.33
False Starts	0.84	1.19
Reformulation	2.31	2.69
Repetition	7.03	5.91
Replacements	1.03	1.40
Complexity	1.70	0.32
Accuracy	0.61	0.19
Number of words	290.41	118.01

Note: $n = 16$

In this study, learners spoke at a rate of 127.45 words per minute, on average. Filled pauses, which were characterized by the number of meaningless words (e.g. "um," "er") both at mid-clause and clause ends, were produced at 13.19 pauses per learner. For each narrative, learners produced 16.91 seconds of silent pauses. For the

measures of repair disfluencies, repetition, which is the number of immediate and verbatim repetitions of a word or phrase, occurred the most in each narrative (7.03 repetitions per learner). Reformulation came next (2.31 reformulations per learner), indicating phrases or clauses that are repeated with some modification either to syntax, morphology, or word order. The number of replacements (1.03 replacements per learner) showed the lexical items that are substituted for another, while false starts, which are utterances that are abandoned before completion, occurred the least in each narrative (0.84 false starts per learner).

As for syntactic complexity, the mean for each learner is 1.70, which indicated the ratio of the total number of clauses to the total number of A-S units (Foster et al., 2000). Concerning accuracy, generalized measures were used. The mean for accuracy, which is the percentage of clauses that did not contain any errors relating to syntax and morphology, is 0.61 per learner.

In general, the fluency scores obtained in this study are lower than those reported in other research which also utilized identical measures to analyse learners' language elicited by narrative tasks (Foster & Skehan, 1996; Skehan & Foster, 1997, 1999). Figures on complexity are higher than those obtained by similar measures in comparable research (Foster & Skehan, 1996; Skehan & Foster, 1997, 1999), whereas accuracy scores are similar to those reported in Foster and Skehan (1996) and Skehan and Foster (1997), but lower than Skehan and Foster (1999), which shows that this type of tasks have some predictable performance characteristics.

4.2.2 Means and Standard Deviations under Four Conditions

Table 7 shows the comparable figures for the 4 conditions under which the tasks were done. Provided in the table are the means and standard deviations for each of the four conditions during task completion. These conditions are 1) *planning, pictures*, 2) *no planning, pictures*, 3) *planning, videos*, and 4) *no planning, videos*.

Table 7

Performance as a Function of Planning and Prompt Conditions

Variable	Planning				No Planning			
	Pictures		Videos		Pictures		Videos	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Speed	152.13	30.02	135.79	19.79	110.36	38.47	116.37	32.4
Pauses (filled)*	11.50	19.11	10.17	17.21	9.63	10.16	19.20	17.45
Pauses (silence)*	15.44	6.93	23.08	8.21	10.25	3.56	19.70	5.02
False Starts*	0.88	1.36	0.83	1.60	0.38	0.74	1.20	1.14
Reformulation*	2.13	1.96	2.67	3.72	1.50	1.60	2.90	3.35
Repetition*	7.25	5.26	5.17	3.19	5.13	3.52	9.50	8.44
Replacements*	0.75	0.89	1.83	2.32	0.50	0.76	1.20	1.40
Complexity	1.59	0.17	1.53	0.28	1.82	0.41	1.79	0.34
Accuracy	0.75	0.10	0.54	0.10	0.54	0.23	0.59	0.22
Number of words	262.25	113.79	367.00	93.06	172.25	64.42	361.50	84.94

*items which indicate higher fluency when scores are low

Learners performing the task under the condition *planning, pictures* produced the highest mean rate of speech (152.13 words per minute), which is followed by *planning, videos* (135.79 words per minute), *no planning, videos* (116.37 words per minute) and *no planning, pictures* (110.36 words per minute). The higher the scores for speed, the higher the rate of fluency of speech.

For the other measures of fluency (both types of pauses and repair disfluencies), it should be noted that the lower the scores are, the higher the scores for fluency of speech. Learners performing the task under the condition *no planning, pictures* produced the least pauses (both filled and silent pauses) and the least repair disfluencies (false starts, reformulation, repetition, replacements). *No planning, videos* elicited the most filled pauses, whereas *planning, videos* elicited the most silent pauses.

In ascending order, the means for filled pauses are: *no planning, pictures* (9.63 pauses per learner) < *planning, videos* (10.17 pauses per learner) < *planning, pictures* (11.50 pauses per learner) < *no planning, videos* (19.20 pauses per learner). For silent pauses, the means are *no planning, pictures* (10.25 pauses per learner) < *planning, pictures* (15.44 pauses per learner) < *no planning, videos* (19.20 pauses per learner) < *planning, videos* (23.08 pauses per learner) in ascending order.

For repair disfluencies, the condition *no planning, pictures* elicited the fewest of these (false starts: 0.38 per learner; reformulation: 1.50 per learner; repetition: 5.13 per learner; replacements: 0.50 per learner). *No planning, videos* elicited the highest means for false starts (1.20 per learner), reformulation (2.90 per learner) and repetition (9.50 per learner). The highest mean for replacements was produced by learners performing under *planning, videos* (1.83 per learner).

For complexity, *no planning, pictures* elicited the highest mean, whereas *planning, videos* elicited the smallest mean. In ascending order, the means for

complexity under the 4 conditions are: *planning, videos* (1.53 per learner) < *planning, pictures* (1.59 per learner) < *no planning, videos* (1.79 per learner) < *no planning, pictures* (1.82 per learner).

For accuracy, learners performing the task under *planning, pictures* produced the most accurate speech, whereas both of the conditions *planning, videos* and *no planning, pictures* produced the least accurate speech. The means for accuracy under the 4 conditions, in ascending order, are: *planning, videos* and *no planning, pictures* (0.54 per learner) < *no planning, videos* (0.59 per learner) < *planning, pictures* (0.75 per learner).

4.3 Correlational Statistics

4.3.1 Intercorrelations of Dependent Variables

Table 8 shows the correlation of the nine dependent variables with one another. Some of the correlations are significant beyond the 0.05 level, others beyond the 0.01 level. The results show that only some of the measures of fluency, including silent pauses, repetition and replacements, intercorrelated significantly with various other fluency measures. These significant correlations include speed, filled pauses and unfilled pauses.

Silent pauses correlated significantly with speed (0.42), filled pauses (0.50), and repair disfluencies such as repetition (0.49) and replacements (0.57). Other than

intercorrelating with silent pauses, filled pauses also correlated significantly with repetition (0.76) and replacements (0.37). The other domains of fluency, including

Table 8

Intercorrelations of Dependent Variables

	Speed	Pauses (filled)	Pauses (silence)	False Starts	Reformulation	Repetition	Replacements	Complexity	Accuracy
Speed	-	0.19	0.42*	0.01	-0.08	0.00	0.27	-0.34	0.33
Pauses (filled)		-	0.50**	-0.07	-0.12	0.76**	0.37*	0.26	-0.06
Pauses (silence)			-	0.13	0.09	0.49**	0.57**	-0.13	0.02
False Starts				-	0.05	0.10	0.16	-0.17	-0.07
Reformulation					-	-0.06	0.19	-0.19	-0.03
Repetition						-	0.31	0.14	-0.05
Replacements							-	0.08	-0.18
Complexity								-	-0.20
Accuracy									-

n = 16

*Correlation is significant at the 0.05 level (2-tailed).

**Correlation is significant at the 0.01 level (2-tailed).

false starts and reformulation, produced no significant correlation against any other measures.

On the other hand, the set of fluency measures generated no significant correlations between the two measures of form, complexity and accuracy. The level of correlation suggests that these two domains are independent of each other for the current tasks. Moreover, there is no significant correlation between these two aspects of form either, as no significance was reached in their negative correlation of –0.20. This negative relationship indicates that higher accuracy in language production can be achieved at the expense of complexity, and vice versa.

4.3.2 Two-way ANOVAs for Fluency, Complexity and Accuracy

The results described above show that although a number of dependent variables were used in the study, using a two-way ANOVA is more appropriate as the main tool for analysis, compared to a multivariate ANOVA. This is because the structure of the dependent measures shows that fluency, complexity and accuracy are three separate factors.

Because "unfilled pauses" generates the highest correlations with other fluency measures, it is used to represent this domain. The results of the two-way ANOVA are presented in Table 9. The means and standard deviations from Table 6 from the previous section will also be referred to, in order to justify for the significance found in Table 9.

Table 9 shows that there is a significant effect for fluency (as indexed by unfilled pauses), in terms of the individual variables of *prompt* (significant at 0.000) and *planning* (significant at 0.054). There is also a significant effect for complexity for *prompt* conditions (significant at 0.051), and a significant effect for accuracy under the combined effects of *prompt* and *planning* (significant at 0.060).

For fluency, there is a marked effect for different tasks. This is also shown in the marked difference between the means of different tasks on fluency measures (see Table 6). *No planning, videos* elicited double the number of filled pauses of *no planning, pictures* (19.20 and 9.63 respectively), and four times the number of reformulations (1.20 and 0.38 respectively). On the other hand, *no planning, pictures* elicited two times fewer silent pauses than *planning, videos* (10.25 and 23.08 respectively), and three times fewer replacements (0.50 and 1.83 respectively).

Pre-task planning did not elicit remarkably fluent speech (see Table 6). Regardless of whether participants looked at pictures or videos, there was a high rate of disfluency markers, indicating participants' high involvement in repair-oriented behaviour. Interestingly, this did not inhibit their rate of speech production. Instead, pre-task planning increased the rate of speech to a remarkable level, as shown in both *planning, pictures* and *planning, videos*. However, it should be noted that prompt and planning only have individual effects on fluency.

Table 9

Two-way ANOVAs for Pauses (silence), Complexity and Accuracy

Source of variance	Sum of squares	*df*	Mean square	*F*	Significance of *F*
Pauses (silence)					
Prompt	566	1	566	16	0.000
Planning	142	1	142	4.03	0.054
Prompt * Planning	6.30	1	6.30	0.18	0.676
Complexity					
Prompt	0.02	1	0.02	0.16	0.692
Planning	0.46	1	0.46	4.60	0.051
Prompt * Planning	0.00	1	0.00	0.02	0.880
Accuracy					
Prompt	0.05	1	0.05	1.46	0.236
Planning	0.06	1	0.06	1.68	0.206
Prompt * Planning	0.13	1	0.13	3.84	0.060

* = interaction between variables

There is also a main effect for complexity under the planning conditions (significant at 0.051). This is justified by referring to the descriptive statistics in Table 6. The means for complexity in Table 6 shows that non-planners produced markedly

more complex speech than planners, regardless of prompt (*planning, pictures:* 1.59; *planning, videos:* 1.53; *no planning, pictures:* 1.82; *no planning, videos:* 1.79).

The picture for accuracy is more complicated. There are no main effects, and neither planning or prompt in isolation exerts significant effects on this dimension. However, there is a significant effect under the combined impact of both planning and prompt (significant at 0.060). Again, the means for accuracy from Table 6 are used as reference for this significant dimension, which is *planning, pictures* (0.75). This condition contrasted with the other three conditions, which are *planning, videos* (0.54), and *no planning, pictures* (0.54), and *no planning, videos* (0.59). Interestingly for the three key aspects of speech production, accuracy is the only dimension under which each of the *planning* and *no planning* conditions have the same mean (0.54).

4.4 Summary of Data Elicited in Language Production

To sum up briefly, pre-task planning has an influence on all three aspects of speech, though maybe not entirely positive. First of all, pre-task planning produces mixed results in fluency. Pre-task planners performing the task on pictures produced the most fluent speech in terms of speed, followed by pre-task planners performing on videos.

However, planners on videos produced a significant amount of silent pauses and replacements, while planners on pictures produced an average amount. Compared with planners, non-planners on pictures produced the fewest number of pauses and

repair disfluencies. The highest number of filled pauses and false starts were produced by non-planners on videos. The type of prompts also has an effect on fluency. Learners performing the task on pictures produced more fluent speech, in terms of a higher rate of speech, and a small amount of pauses and repair disfluencies. This is especially true for non-planners performing on pictures. On the contrary, regardless of whether learners were given the opportunity to plan or not, those performing the task on videos always produced the highest rate of pauses and repair disfluencies.

Second, results show that pre-task planning does not have a positive effect on the complexity of learners' speech. Non-planners produce language which is more complex than planners, implying that pre-task planning might have a detrimental effect on complexity. Last of all, the combined effects of planning and prompt have an influence on accuracy. However, it is hard to distinguish a pattern between pre-task planning and prompt type. Pre-task planners performing the task on pictures produce the most accurate language, followed by non-planners on videos. The same level of accuracy occur in the language production under the conditions *planning, videos* and *non-planning, pictures*. The reasons as to the mixed results of accuracy will be speculated upon in the next chapter.

4.5 Learners' Perception of Task Performance

4.5.1 Learners' Perception of Task Difficulty and Task Enjoyment

In this section, learners' perception of task difficulty and task enjoyment is

examined. In a retrospective task difficulty questionnaire (see Appendix D), learners were asked to rate these two items, task difficulty and task enjoyment, in a 5-point Likert scale. In terms of task difficulty, 1 represents that the task is least difficult, whereas 5 represents that the task is most difficult. In terms of task enjoyment, 1 represents that the task is least enjoyable, whereas 5 represents that the task is most enjoyable.

4.5.1.1 Descriptive Statistics: Learners' Perception of Task Difficulty and Task Enjoyment

Table 10 shows the means and standard deviations of pre-task planning and prompt type, when they interact with task enjoyment and task difficulty. Pre-task planning has a positive effect on task enjoyment. Learners clearly find the task more enjoyable when they have the opportunity to plan (mean = 3.86) than not (mean = 3.28). Pre-task planning also eases the difficulty of the task for planners (mean = 2.86), as non-planners find the task more difficult (mean = 3.22).

The type of prompt also produces a difference in learners' perception of task enjoyment. Learners find video prompts markedly more enjoyable than picture prompts (pictures: mean = 2.93; videos: mean = 4.13). However, performing the task with more enjoyment does not seem to have an effect on task difficulty. Dealing with either pictures or videos does not ease the task in any way (pictures: mean = 3.06; videos: mean = 3.06).

Table 10

Learners' perception of task enjoyment and difficulty

Task dimension and version		Task difficulty		Task enjoyment	
		3.5 minutes	*0.5 minutes*	*3.5 minutes*	*0.5 minutes*
Planning	***M***	2.86	3.22	3.86	3.28
	SD	0.66	0.88	0.66	0.96
	n	14	18	14	18
		Pictures	*Videos*	*Pictures*	*Videos*
Prompts	***M***	3.06	3.06	2.93	4.13
	SD	0.93	0.68	0.85	0.34
	n	16	16	16	16

4.5.1.2 Correlational Statistics: Learners' Perception of Task Enjoyment

A two-way ANOVA was used for analysis to look for significant interactions between the independent variables of pre-task planning and prompt, with learners' perception of task enjoyment. The results are shown in Table 11.

Table 11 shows a further analysis of the interaction between planning and prompt type with task enjoyment. There are both main effects on task enjoyment, and a two-way interaction which indicates that planning and prompt do have an influence in

certain combinations.

Concerning pre-task planning, Table 11 shows that pre-task planning has a significant effect on task enjoyment (significant at 0.000). Learners enjoyed the task more when they are allowed to plan. This is true for all participants, regardless of whether they were performing the tasks on pictures or videos.

Table 11

Two-way ANOVA for task enjoyment

Source of variance	**Sum of squares**	***df***	**Mean square**	***F***	**Significance of *F***
Main effects					
Planning	4.12	1	4.12	15.99	0.000
Prompt	11.70	1	11.70	45.44	0.000
Planning* Prompt	1.21	1	1.21	4.71	0.039

* = interaction between variables

Concerning prompt type, there is again a significant effect on task enjoyment (significant at 0.000). Table 11 shows that regardless of proficiency level or the availability of pre-task planning, all learners agree that videos are much more enjoyable as a tool for story-retelling than pictures. This goes in line with previous discussions above, that the structure, amount of interaction and audio-visual attraction

of videos can be great stimulants of interest, as compared with picture sequences. There is also a significant effect on task enjoyment when planning and prompt are combined together (significant at 0.039).

In general, learners find that the opportunity to plan before the task enhances their enjoyment of it, while at the same time they also perceive that it also lessens the cognitive demand of the task. The prompt type of the task also makes a difference to the level of enjoyment that learners experience while performing the task. Learners find video prompts much more enjoyable than picture sequences although this does not affect their perception of task difficulty.

4.5.2 Learners' Planning Strategies and Difficulties Encountered While Speaking

This section examines learner strategies in terms of different proficiency levels. Advanced and intermediate learners were asked to complete a retrospective questionnaire, which focuses on three areas: 1) whether learners find pre-task planning useful or adequate, 2) what they did when they planned, and 3) what would help them improve their task performance. The questionnaire was designed to tap learners' reactions to their story-telling experience without drawing their attention to particular conditions, thus it is in self-report style (see Appendix D).

The first two items are designed exclusively for learners who had the opportunity to plan before the task.

1. Is planning time useful or adequate?

Table 12 shows that all intermediate learners find planning time useful (100% of all intermediate planners), while slightly fewer advanced learners find it useful (88% of advanced planners). The reverse pattern is true for adequacy of planning time. More advanced learners think that 3.5 minutes of planning time is adequate (88% of advanced planners), whereas only 50% of intermediate planners find it adequate. In general, learners agree that planning time is useful, but only slightly more than half of them think that the planning time provided is adequate.

Table 12

Learners' Perception of the Usefulness and Adequacy of Planning Time

Planning time	**Advanced learners**	**Intermediate Learners**	**Total**
Usefulness	88%	100%	94%
Adequacy	88%	50%	69%

2. How did you spend your planning time?

Learners responded to this question item in self-report style. The results are shown in Table 13, which indicates that most learners, regardless of their proficiency level, spent their planning time on organization (34% of all planners). This is followed by vocabulary (24% of all planners). Recalling the details of the story and

comprehension are also important parts of pre-task planning (17% and 14% of all planners respectively). A small proportion of both advanced and intermediate learners also spent some time on listener strategies (6% of all planners). Interestingly, none of the planners spent any time on grammar.

Table 13

Learners' Perception of How to Spend Their Planning Time

Categories	Response	Advanced learners	Intermediate learners	Total
Language	Grammar	0%	0%	0%
	Vocabulary	23%	25%	24%
Plot	Comprehension	15%	13%	14%
	Organization	38%	31%	34%
Presentation skills	Listener strategies	8%	6%	6%
Rehearsal	Memorization	0%	6%	3%
	Recalling details	15%	19%	17%
	Summarizing	0%	0%	0%

The following question items were answered by both planners and non-planners.

3. What would help improve task performance?

The categories shown in Table 14 emerge from the data, according to the responses of the learners. The results show that the majority of learners, especially non-planners, thought that the opportunity to plan before the task would improve their performance (29% of advanced learners; 57% of intermediate learners). This is especially true for intermediate learners, whose percentage for this question is double that of advanced learners. 21% of advanced learners also thought that having their notes or the picture prompts in front of them when they perform the task would be helpful. Group discussion before re-telling the story was also considered to help with task performance by 16% of the advanced learners and 10% of intermediate learners.

4% of advanced learners and 10% of the intermediate learners thought that having more dialogue in the picture prompts would help, while another 4% of advanced learners and 10% of intermediate learners thought that being nervous when they re-told the story prevented them from performing well.

Other response items which involve a smaller proportion of learners, include the providing more vocabulary (8% of advanced learners; 0% of intermediate learners), using videos instead of pictures as the narrative prompt (4% of advanced learners; 5% of intermediate learners), having English subtitles in the video snippet (4% of advanced learners; 5% of intermediate learners), and familiarity with the task (4% of advanced learners; 5% of intermediate learners). Interestingly too, 4% of advanced

learners expressed that self-improvement in English would help i.e. increasing their own exposure to English, while none of the intermediate learners expressed a similar strategy. None of the participants think that grammar would help much.

Table 14

Learners' Perception of What Would Help Them Improve Task Performance

Categories	Response	Advanced Learners	Intermediate Learners	Total
Emotional	Nervous	4%	10%	6%
Non-planners	Planning	29%	57%	42%
Pictures	More dialogue	4%	10%	6%
	Use videos	4%	5%	4%
Videos	Eng subtitles	4%	5%	4%
Interaction	Group discuss	16%	0%	8%
Language	Lack vocabulary	8%	0%	4%
	Grammar	0%	0%	0%
Self-improvement	More exposure to Eng	4%	0%	2%
Task	Familiarity	4%	5%	4%
Immediacy	Allowed to read notes / pictures when narrating	21%	10%	16%

4.5.3 The Benefits of Different Prompt Types

Participants were asked if they had any additional comments on the task in the retrospective questionnaire. Many of them expressed their opinion on the prompts that were used. Their response is categorized in the following section.

Table 15 shows that many learners perceived that videos offer more benefits for task performance. Learners felt that they benefited most from the vocabulary provided in the videos (23% of advanced learners; 33% of intermediate learners). Learners also found the language in the videos easy to understand (6% of advanced learners; 33% of intermediate learners). Again this was especially true for intermediate learners, as the percentage of intermediate learners is more than five times the amount of advanced learners in agreeing to this point. Moreover, 11% of advanced and 11% of intermediate learners also thought that the organization of video prompts is easy to follow.

Other benefits of video prompts include audio-visual stimulation (6% of advanced learners; 11% of intermediate learners), comprehension of the topic (12% of advanced learners; 0% of intermediate learners), having an interesting topic (6% of advanced learners; 11% of intermediate learners), have an appropriate length (9% of advanced learners; 8% of intermediate learners), and easy to recall details, especially when note-taking was allowed (6% of advanced learners; 0% of intermediate learners).

Table 15

Learners' Perception of the Benefits of Different Prompt Types

Categories	Response	Advanced learners		Intermediate learners		Total	
		Pictures	Videos	Pictures	Videos	Pictures	Videos
Stimulation	Audio-visual	0%	6%	0%	11%	0%	8%
Language	Provide vocabulary / phrases	0%	23%	0%	33%	0%	26%
	Simple	0%	6%	0%	33%	0%	15%
Organization	Follow story	6%	11%	0%	11%	4%	7%
Plot	Comprehension	0%	12%	0%	0%	0%	8%
	Interesting topic	0%	6%	0%	11%	0%	8%
	Room for imagination	6%	0%	0%	0%	4%	0%
	Appropriate length	9%	9%	0%	0%	8%	8%
Note-taking	Recalling details	0%	6%	0%	0%	0%	4%

Some learners thought that picture prompts were also beneficial to their task performance in some ways. 9% of advanced learners found that the picture sequence was appropriate in length. 6% of advanced learners thought that the organization of picture prompts is easy to follow, while another 6% of advanced learners found that pictures offer them room for imagination. Interestingly, no intermediate learners thought that picture prompts are beneficial in any way.

4.6 Summary of Learners' Perception of Tasks

To sum up learners' perspective of the tasks, both advanced and intermediate learners thought that pre-task planning benefitted their task performance. In general, learners of both proficiency levels focused on organizing the story, retrieving vocabulary from their mental lexicon and comprehending the story when they plan. There does not seem to be a major difference between the strategies they used.

Concerning what learners thought would help improve their task performance, a significant proportion of non-planners, especially intermediate learners, stated that pre-task planning would help. In general learners thought that having their notes or the picture prompt in front of them when they re-told the story would help ease the difficulty of the task. Advanced learners especially thought that having group discussions before re-telling the story would help their comprehension of the story, and would also allow them to benefit from the grammatical and lexical sources of other learners.

All learners agreed that video snippets were more beneficial as narrative prompts. This is especially because videos provide lots of vocabulary and phrases in dialogue that learners can recycle in their narrative tasks. A lot of intermediate learners also thought that the language in the videos was simple, thus enabling easy comprehension. Other benefits of video prompts include possessing audio-visual stimulation, having an interesting topic, providing a clear structure and being appropriate in length. Advanced learners also considered picture prompts as beneficial to the task with due to a clear structure, room for imagination and appropriate length. Interestingly, none of the intermediate learners thought that picture prompts were beneficial in any way. The implications of these results will be interpreted and presented in the next section.

CHAPTER FIVE

DISCUSSION

5.1 Introduction

This study sought to examine the different effects of pre-task planning on fluency, accuracy and complexity of Hong Kong university students' oral production in a narrative task. This section begins by addressing the research questions and comparing the statistical findings from this study with those from previous research. Learners' perceptions as derived from the questionnaires are also considered in the interpretation of the results.

Four research questions were proposed earlier in this study. The following discussion starts with a restatement of each research question (see Literature Review). A summary of the findings is briefly given, followed by justifications drawn from previous literature and *post-hoc* interpretation.

5.2 Research Question One

What effect does pre-task planning, in addition to the use of picture prompts and video prompts, have on the *fluency* of L2 learners' production in an oral narrative task?

5.2.1 Summary of Results on Fluency

The results show that the fluency of L2 learners is not totally enhanced by pre-task planning. The measures of fluency in this study can be categorized into speed, pauses (filled and unfilled) and repair disfluency (false start, repetition, reformulation, replacement). The only measure of fluency which shows an enhancement in planning conditions is speed, where the planning conditions elicit 152.13 and 135.79 words per minute on average respectively for pictures and videos, as opposed to 110.36 and 116.37 words per minute respectively for unplanned conditions. Planners on average speak markedly faster (hence more fluently) than non-planners.

The investigation on pauses shows mixed results. Non-planners produced fewer pauses during their story-retelling, while planners produced more filled pauses than non-planners. For filled pauses (the number of meaningless words such as "um" and "er" both at mid-clause and clause ends), planners for both picture and video prompts produced fewer pauses, together with non-planners on picture prompts (11.50, 10.17 and 9.63 filled pauses are elicited from each task respectively). Non-planners for videos produced the most pauses (19.20 filled pauses).

The results for unfilled pauses (the length of silent pauses at both mid-clause and clause ends) is even more confusing, with non-planners for pictures producing the fewest pauses on average (10.25 seconds of silent pauses), planners for pictures producing more pauses (15.44 seconds of silent pauses) than non-planners for videos (19.70 seconds of silent pauses), with planners on videos pausing the most during the

narrative (23.08 seconds of silent pauses). In other words, planning does not reduce learners' hesitations during their speech, and in some cases even cause more pausing. However, it should be noted that pre-task planning is statistically significant at 0.000 for fluency.

The results on repair disfluency are consistent with the mixed results for pauses. (For easy reference, false starts are utterances that are abandoned before completion. Repetition stands for immediate and verbatim repetitions of a word or phrase. Reformulation is a phrase or clause that is repeated with some modification either to syntax, morphology, or word order. Replacement stands for a lexical item that is substituted for another).

All measures of repair disfluency, except for replacement, reveal that both non-planners for pictures and videos produce the fewest and most disfluencies respectively (false starts: mean = 0.38 and 1.20; repetition: mean = 5.13 and 9.15; reformulation: mean = 1.50 and 2.90), while planners for pictures and videos fare in between the two (false starts: mean = 0.88 and 0.83; repetition: mean = 5.17 and 7.25; reformulation: mean = 2.13 and 2.67). For replacement, the findings show even less consistency with the prediction that pre-task planning reduces the amount of repair disfluencies in language learners as non-planners for pictures produce the fewest replacements (mean = 0.50). However, there is only a small difference with planners on pictures (mean = 0.75). Non-planners on videos produce fewer replacements than planners on videos (mean = 1.20 and 1.83 respectively). In short, planning conditions

do not seem to give planners a great advantage over non-planners in terms of repair disfluencies.

Generally speaking, speed is the only fluency measure which shows consistency with previous research (Bygate, 1996; Crookes, 1989; Foster & Skehan 1996, 1997; Ortega 1999; Skehan & Foster 1997), and also conforms to the prediction of this study that pre-task planning improves fluency. There is a significant difference between planners and non-planners of the same prompt, with perhaps pictures being the more indicative prompt, as planners who performed the task on pictures produced the most speed. However, for pauses and repair disfluencies, which indicate greater fluency when means are small, show that videos are more indicative of the planning condition, as non-planners doing videos often produce more pauses and disfluencies than planners. The situation is different for picture prompts, which show that non-planners of picture prompts fare better than planners.

5.2.2 Explanations for Results on Fluency

What explanation is there for the difference in results for fluency between this study and previous research? Tavakloli and Skehan (2005) concluded that "fluency is a complex construct" (p. 267). Despite the traditional claim that pre-task planning has a positive effect on fluency, there is a need to distinguish between the various measures of fluency. After examining fluency in more detail, it is found that not all measures of fluency conform to this claim. Foster and Skehan (1996) reported that

planners pause less frequently and spent less time in total silence than non-planners. Skehan and Foster (1997) found the same for total pauses. Wendel (1997, as cited in Yuan and Ellis, 2003) found that planners produce more syllables per minute, and also show a lower mean length of pauses than non-planners. Ortega (1999) found a faster speech rate in planners.

Therefore, it appears that pre-task planning does not have a positive effect on *every* measure of fluency; the main effects are on speech rate and pause length. However, for repair disfluencies, Tavakoli and Skehan (2005) reported that interestingly, "each of these measures is higher for the planned condition i.e. planners seem to be more likely to engage in modification of speech on-line" (p. 268), though none of these measures were statistically significant. They drew the conclusion that there is a separation between repair fluency and breakdown fluency/speed on the other, suggesting that "a concern to modify utterances online is somewhat distinct from a capacity to organize speech in real time" (p. 268). This suggests the need to examine what strategies learners use when they monitor their speech on-line, which at this point have received little attention in the research on planning.

It is also important to note that different prompts were used in this study. As opposed to previous studies which traditionally use picture prompts, this study used both picture sequences and video snippets. Concerning pauses, picture prompts do not seem to be indicative of the positive effects of pre-task planning, while video prompts do confirm findings from previous studies: that non-planners produce more pauses

(Foster & Skehan, 1996, 1997; Ortega, 1999; Skehan & Foster, 1997). Similarly for repair disfluencies, picture prompts do not confirm the prediction of this study, while video prompts do. To account for this situation, the operationalization of narrative tasks with picture prompts and video prompts must be examined.

An obvious difference between tasks with picture prompts and those with video prompts, is that while non-planners performing the task with picture prompts have only 0.5 minutes to look at the pictures *and* prepare for their speech, the non-planner group performing the task with video prompts have 7 minutes to watch the video *and* 0.5 minutes to prepare for their speech. Similarly for the planner group, learners performing the task on pictures have only 3.5 minutes to plan before the task, whereas learners performing the task on videos have 7 minutes to watch the video *and* 3.5 minutes to plan before the task. Thus when the tasks are organized in ascending order according to the amount of time learners have to plan for their speech before the task, the order should be no *planning, pictures* (0.5 minutes) > *planning, pictures* (3.5 minutes) > *no planning, videos* (7.5 minutes) > *planning, videos* (10.5 minutes). Because learners performing the task on video prompts confirmed the prediction that pre-task planning benefits fluency, this might suggest that the difference between the language production of planners and non-planners is clear when there is a significant difference in the amount of pre-task planning time.

Taken broadly, these results indicated support for the claim that pre-task planning leads to an increase in the main dimension of fluency, as confirmed by previous

research (Foster & Skehan, 1996, 1997; Skehan & Foster, 1997, Ortega, 1999). However, the positive effect of pre-task planning on complexity is not so clear. The syntactic complexity of learners' language production in this study is discussed below.

5.3 Research Question Two

What effect does pre-task planning, in addition to the use of picture prompts and video prompts, have on the *complexity* of L2 learners' production in an oral narrative task?

5.3.1 Summary of Results on Complexity

The results show that non-planners produce more syntactically complex narratives than planners, though the difference is small. In descending order, the means for complexity are 1.82 (non-planners on pictures) > 1.79 (non-planners on videos) > 1.59 (planners on pictures) > 1.53 (planners on videos). A main effect has been shown for complexity under planning conditions (significant at 0.051). However, the results are somewhat surprising. Previous studies found complexity enhanced with pre-task planning. Crookes (1989) reported that 10 minutes of planning time results in greater structural complexity and lexical variety. Foster and Skehan (1996) reported that planners produce significantly more subordination than non-planners. Mehnert (1998) found a positive effect for 10-minutes planners, whereas a similar level of complexity is obtained by 1-minute and 5-minute planners and non-planners. Yuan

and Ellis (2003) and Tavakoli and Skehan (2005) also reported greater complexity with pre-task planning, though the effect size is small for the latter study.

This also contrasted with Skehan and Foster (1999), who used *Mr. Bean* videos as the narrative prompt and found that the pressing need to keep up with video-watching in real time and then tell the story (with no planning time given) results in less complex language being used by the participants. These studies showed that complexity is enhanced when learners have a reasonable length of time to plan.

5.3.2 Explanations for Results on Complexity

What explanation is there for the difference in results for complexity between this study and previous research? One of the most obvious reasons concerns the limitations of complexity measures. Crookes (1989) measured syntactic complexity and lexical variety. Skehan and Foster (1997) examined the index of subordination, while Ortega (1999) looked at the number of words per utterance. Yuan and Ellis (2003) explored syntactic complexity and syntactic variety. The generalization can be drawn that different researchers look at various aspects of complexity, which sometimes overlap with other studies.

Other than different operationalizations, the measures of complexity used in these studies do not include *all* complexity measures. Ellis (2005) points out that no previous studies distinguish between propositional complexity (i.e. the content of the

learners' messages) and formal complexity (i.e. the actual language used). This study uses only a generalized measure of complexity. Obviously, further research is needed to find out whether different complexity measures make a difference to the results.

There are also different models of task-based performance and learning. Skehan's (1998) cognitive approach to tasks distinguishes between three aspects of production: fluency, accuracy and complexity, and suggests that learners vary in the extent to which they emphasize each aspect. Assuming that learners possess a limited processing capacity, the trade-off effect will occur among these three aspects, which results in learners making decisions on how to allocate their attentional resources (see Literature Review). Skehan's (1998) model is also the general approach to tasks for many studies on planning (Iwashita et al., 2001; Wigglesworth, 1997; Yuan & Ellis 2003).

However, there is another approach to tasks, suggested by Robinson (2001). His view of task difficulty is different from Skehan's (1998), in that he states that the more cognitively demanding the task is, the more accurate and complex the learner gets, whereas cognitively less demanding tasks promote fluency only. He argues that learners are like native speakers, and have the capacity to attend to more than one aspect of language at the same time.

In his theory, task complexity is determined by two sets of features. The first of these is the "resource-directing" factor, which includes the number of task elements, the reasoning demands within the task, and whether the immediacy of information is

provided. Next is the "resource-depleting" factor, which consists of pre-task planning time, the number of tasks the learner has to complete, and prior knowledge to the task. By manipulating these factors, the cognitive demand required for task performance will vary, which in turn leads to the variation in quality of the language produced.

According to Robinson (1995, 2001), easy tasks are those which involve resource-depleting factors, such as having contextual support, no reasoning demands and allowing for learners to plan for the task. This type of task promotes fluency at the expense of accuracy and complexity. On the contrary, harder tasks are those which involve resource-directing factors. In short, these tasks offer no contextual support, have lots of reasoning demands, and offer no pre-task planning time. In such cases, Robinson (1995, 2001) claims that learners will "rise to the challenge" and draw on a greater range of syntactic resources than when they are given an easier task to do, thus enhancing both complexity and accuracy but not fluency.

This is a possible explanation for the clear difference between the complexity levels of non-planners and pre-task planners in this study. If Robinson's (2001) model is followed, then non-planners who face "resource-directing" factors are forced to draw on a greater range of syntactic resources, thus producing narratives that are more complex than pre-task planners. Whereas pre-task planners, on the other hand, face "resource-depleting" factors, which pose a detriment to their complexity level. In this interpretation, the opportunity for pre-task planning becomes an obstacle to achieving more complex and accurate language – in contrast to Skehan's (1998) claim that

pre-task planning is beneficial to complexity, if not always accuracy. However, it must be noted that the majority of studies on pre-task planning have findings which supported Skehan (1998)'s framework of task performance. (See Literature Review).

It is noteworthy that the results for complexity show that there is a clear-cut difference between planners and non-planners. In other words, planners spoke with less variety in syntax and lexis, and possibly also with simpler sentence structure. On the other hand, non-planners produce sentences which are more varied and complex in structure, and possibly also more varied in vocabulary.

This can be explained by citing learners' self-report of how they utilized their pre-task planning time, which shows that planners, regardless of proficiency level, showed concern for the comprehension of the listener during task performance (8% of advanced learners; 6% of intermediate learners). Ortega (2005) reported that there is a "natural divergence in task approach" (p. 91) between some learners with an overall tendency towards communication and some who are more inclined towards accuracy. She characterizes learners who "spoke of successful communication in real-world situations as self-regulated through the features of context and interaction" (p. 91) as communication-oriented. These learners do not worry about making mistakes and even view errors as a gradual process of second language learning. Accuracy is not a realistic goal for them.

On the other hand, learners "who have a lesser concern for the impact of their performance on the listener" and even think that speaking correctly is "more

important than the quality of the content of the message" (Ortega, 2005, p. 92) tend to see L2 learning as a "prolonged effort to reach 'a hundred percent correctness" (p. 92). These learners are labeled accuracy-oriented. Although Ortega (2005) distinguished this difference in learner orientation, this can also be applied to the difference between planners and non-planners. Planners who have the opportunity to plan seem to fall into the category of communication-oriented learners who are concerned for their performance on the impact of the listener, while non-planners are more concerned for the quality (i.e. grammatical accuracy) of their speech instead. Because simpler vocabulary and sentences are more listener-friendly, planners unconsciously aimed for this in the story when they considered listener strategies, while non-planners aimed for greater syntactic complexity instead of communicative strategies.

Other than the communicative goal of learners, a number of studies indicated that there might be a possible trade-off effect between complexity and accuracy (Crookes, 1989; Foster & Skehan, 1996; Mehnert, 1998; Skehan & Foster, 1997). Foster and Skehan (1996) reported that guided planners complexify a task at the expense of grammatical accuracy. Skehan and Foster (1997) showed that planners of a decision-making task (inherently unstructured) use planning time to sort out how to express complex ideas to the detriment of accuracy. They claimed that task characteristics play an important role in "channeling the effects of planning towards accuracy and complexity" (p. 48).

Mehnert (1998) also found that the length of time becomes an important factor

when it comes to the allocation of attentional resources. She reported that 1-minute planners give priority to accuracy as compared with non-planners, while 5-minute and 10-minute planners performed at the same overall level of accuracy as 1-minute planners. Concerning complexity, positive effects were only found for 10-minute planners, while planners with 1 minute and 5 minutes performed at the same level of complexity as non-planners. Mehnert's (1998) study showed that learners with limited planning (1 minute and 5 minutes) predispose their attention to more accurate language instead of more complex language, whereas planners with unlimited pre-task planning time (10 minutes) focus on producing more complex language instead of more accurate language. Yuan and Ellis (2003) and Tavakoli and Skehan (2005) similarly reported a trade-off between accuracy and complexity. In answering the next research question, a clearer picture as to whether this is the case for the present study is presented.

5.4 Research Question Three

What effect does pre-task planning, in addition to the use of picture prompts and video prompts, have on the *accuracy* of L2 learners' production in an oral narrative task?

5.4.1 Summary of Results on Accuracy

Pre-task planning has no individual effect on accuracy. There is only a combined

effect of planning and prompt on accuracy. However, the difference is small between the language production of planners and non-planners (significant at 0.060). The most grammatically accurate narratives are produced by planners on pictures (mean = 0.75), followed by non-planners on videos (mean = 0.59). Planners on videos and non-planners on pictures have the same mean (mean = 0.54). This is not surprising, as accuracy is traditionally a construct that produces mixed results in previous research (Ellis, 1987; Foster & Skehan, 1996; Mehnert, 1998; Skehan & Foster, 1997). Ellis (1987) reported that planning has positive effects on only regular past tense forms, but not on the irregular past. Foster and Skehan (1996) stated that while pre-task planning did have an impact on accuracy, the condition has to be unguided planning.

Skehan and Foster (1997) found that accuracy is enhanced with pre-task planning when combined with personal or narrative tasks, but not with decision-making tasks. Mehnert (1998) found that the accuracy of 1-minute planners benefits from pre-task planning, but not with 5-minute or 10-minute planners. Ortega (1999) reported that accuracy is enhanced with pre-task planning on noun-modifiers, while a follow-up study reports that pre-task planning has a positive effect on articles (Ortega, 2005). Skehan and Foster (1999) and Tavakoli and Skehan (2005) found accuracy improved with the combined effects of pre-task planning and structure, but that proficiency level had the greatest effect on accuracy.

These studies concluded that accuracy is influenced by a number of factors, including 1) the type of planning (accuracy is increased with unguided planning), 2)

the type of task (personal or narrative tasks favour accuracy), 3) the grammatical feature explored (when the feature involves a clear rule, accuracy is enhanced), 4) task complexity (a task which is cognitively less demanding favours accuracy), 5) task structure (structured tasks favours accuracy), 6) the length of planning time (limited planning time leads to increased accuracy), and 7) learner proficiency (advanced learners are able to benefit more from planning). In addition to this, it is found that the learners in a language testing situation focus their language production on grammatical accuracy (Iwashita et al. 2001). Does this mean that adding up all these variables in favour of accuracy would guarantee improvement in the accuracy of learners' language production?

Under Skehan's (1998) framework of tasks, Iwashita et al. (2001) tried to predict task difficulty in oral tasks by manipulating certain task characteristics in ways which make the task easier (less cognitively demanding) or more difficult (more cognitively demanding). Iwashita et al. (2001) examined four dimensions, including perspective, immediacy and adequacy and planning time. For perspective, learners were required to tell a story either as it happened to themselves (first person) or someone else (third person). The second dimension is immediacy, which required learners to tell the story with and without the pictures in front of them. The third dimension, adequacy, required learners to tell the story with a complete set of six pictures or with an incomplete set of pictures. The last dimension, planning time, is operationalized in a way similar to the present study. Learners are either performing the task with the

opportunity of pre-task planning (3.5 minutes) or not (0.5 minutes). Each of these dimensions contained a performance condition which is less cognitively demanding and more cognitively demanding.

Iwashita et al. (2001) anticipated that learners performing the task in first person narration, with the complete pictures in front of them and with the opportunity to plan beforehand would produce language which was more accurate, while learners performing the task in third person narration, without the full set of pictures, without the opportunity to look at the pictures as they tell the story and without any planning time would produce language which was less accurate. This study was conducted in a language testing situation, with participants being pre-university students enrolled in a ESL course in Australia, who had TOEFL scores ranging from 427 (low-intermediate) to 670 (advanced).

Iwashita et al. (2001) managed to fulfill at least half of the generalizations mentioned above, which are drawn from previous research. Unguided planning, the use of narrative tasks, low task complexity, limited planning and the language testing condition all favoured accuracy. However, Iwashita et al. (2001) failed to confirm the findings of existing research even with the manipulation of task variables. The only significant finding in their study was the opposite of their hypotheses -- that accuracy increases with the more cognitively demanding condition (under Immediacy) instead of decreases. Similar to Iwashita et al. (2001), the present study manipulated variables comparable to theirs, with the addition of structured prompts and learner proficiency,

which should again favour accuracy. The conclusion that accuracy was not significantly advantaged under pre-task planning conditions in this study showed that accuracy is an unpredictable construct that needs further research on its nature. Greater amounts of planning time might also result in language which is more grammatically accurate.

5.4.2 Explanations for Results on Accuracy

What explanation is there for the mixed results for accuracy between this study and previous research? One of the most obvious reasons is the different operationalizations of accuracy. Previous research studied accuracy using a variety of specific measures, including regular and irregular past tense forms (Ellis, 1987), plurals and verb morphology (Wigglesworth, 1997; Yuen & Ellis, 2003), noun-modifiers and articles (Ortega 1999, 2005). Generalized measures which "though more blunt, capture more variance in performance" (Tavakoli & Skehan, 2005, p. 256) have also been used (Skehan & Foster, 1997, 1999; Tavakoli & Skehan, 2005). Again, the different measures used in various studies make it difficult to compare the results of these studies.

This suggests multiple possibilities for research in accuracy. One of these is that the different grammatical items used to measure accuracy (e.g. past tense forms, verb morphology, articles etc.) might not all be enhanced by pre-task planning. As Ellis (2005) stated, the effect of pre-task planning on accuracy depends on the grammatical

feature explored. Thus future research that explores accuracy must clarify the grammatical feature to be examined; when the feature involves a clear rule, an improvement in accuracy is anticipated.

Other than this, two other possibilities for the lack of pre-task planning effects on accuracy are suggested in previous studies. First, a trade-off effect exists between accuracy and complexity and/or fluency. Studies showed that a trade-off effect possibly involves complexity and accuracy (Crookes, 1989; Foster & Skehan, 1996; Mehnert, 1998; Skehan & Foster, 1997), while a limited number of studies also reported the opposite (Wendel, 1997, as cited in Yuan & Ellis, 2003; Yuan & Ellis, 2003).

Results from Crookes (1989) showed that while there is a strong effect of pre-task planning on fluency and complexity, accuracy remains unchanged. Foster and Skehan (1996) reported that guided planners give themselves a more complex task to do, thus they are less accurate. Skehan and Foster (1997) concluded that planners prioritize either complexity or accuracy, but not both. Mehnert (1998) showed that 1-minute planners devote their attention to accuracy instead of complexity within the limited time, while 10-minute planners do the opposite.

However, there are also studies showing that the trade-off involves accuracy and fluency. Wendel (1997, as cited in Yuan & Ellis, 2003) thought that the type of planning decides whether learners predispose their attention to accuracy or fluency, and that with pre-task planning, learners attend to fluency, whereas the increase in

accuracy is involved with online planning. Yuan and Ellis (2003) also displayed results which confirmed Wendel's (1997, as cited in Yuan & Ellis, 2003) findings. They even suggested a dual trade-off effect, indicating that the primary competition involves fluency and accuracy, while the second trade-off indicates mutual tension between complexity and accuracy.

The discussion on the trade-off effect between the key aspects of speaking essentially points at the limitation of learners' attentional resources. Where learners are willing to allocate more attention to one area, the other two aspects are compromised. However, the results of this study did not find any significance with pre-task planning effects on accuracy. Manipulating task variables according to Skehan's (1998) framework is not enough; the absence of consistent differences in performance suggests that there is a need to define what is meant by various task dimensions and conditions specifically (Iwashita et al., 2001).

Another possible explanation for the lack of accuracy effect in this study might be its sole focus on pre-task planning. Researchers found that pre-task planning has no positive effect on accuracy while on-line planning does (Hulstijn & Hulstijn, 1984; Yuan & Ellis, 2003). Hulstijn and Hulstijn (1984) reported that when learners use their speaking time to monitor their grammatical resources, their production becomes more accurate. A more recent study, Yuan and Ellis (2003), found that pre-task planning does not affect accuracy strongly, which falls in line with the findings of previous studies on planning. On the other hand, the on-line planning group achieved

significantly greater accuracy than both pre-task planning and non-planning groups, whose speaking time is restricted and thus are pressured to perform the task rapidly.

The results of Yuan and Ellis (2003) confirms not only Wendel's (1997, as cited in Yuan & Ellis, 2003) claim that accuracy arises as a result of what learners do during performance, but is also compatible with Skehan's (1998) dual-processing model, which shows that learners' rule-based system requires more time and attentional capacity to access than lexically stored knowledge, and the opportunity to access their rule-based system during performance has a strong positive effect on accuracy.

Similarly, Skehan and Foster (1997) also found that learners may be predisposed to use their planning time to pay attention to organizing and encoding propositional content rather than searching their linguistic repertoire to maximize accuracy. In this study, learners are restricted to speaking for 3 minutes, which might not be enough for learners to re-tell a story with more detail, such as the video prompt utilized in this study.

Thus having limited opportunity to plan on-line might be a possible reason for learners to produce a lack of accuracy, though this does not explain the reason other studies that focused on examining pre-task planning have also found effects on accuracy. Clearly more research is needed to define the construct of accuracy specifically before any definite conclusions can be drawn.

5.5 Summary on the Language Production of Learners

It is found that learners in this study show that with pre-task planning, fluency is enhanced, but not complexity or accuracy. The lack of complexity effects can be explained by the priority of the learners to communicate to the listener, rather than complexifying their narratives. Robinson's (1995, 2001) model of task performance can also explain why there is a lack of positive effects of pre-task planning on complexity and accuracy. In spite of this, previous studies (Foster & Skehan, 1996; Mehnert, 1998; Skehan & Foster, 1997; Wigglesworth, 1997) have shown that pre-task planning *does* benefit both aspects, which is contrary to Robinson's (1995, 2001) claim that pre-task planning enhances fluency at the expense of complexity and accuracy. Thus Robinson's (1995, 2001) model does not apply in this study.

The explanation that remains, therefore, is the trade-off effect between the three aspects: fluency, complexity and accuracy, and also the possibility of limitations for on-line planning. As the results obtained for accuracy does not conform to any consistent pattern, it is impossible to decide on either justification for the present.

5.6 Research Question Four

What is learners' perception of the usefulness of pre-task planning time, task enjoyment or task difficulty? What do learners do when they plan?

5.6.1 Summary of Results on Learners' Perception of Task Performance

Following the discussion on descriptive statistics, the issue of how pre-task

planning influences learners' language production is further examined by analyzing the task difficulty questionnaires. This section focuses on exploring the influence of language expertise on four areas: 1) whether learners find pre-task planning useful or adequate, 2) what they did when they planned, and 3) what would help them improve their task performance. In addition, the benefits of different prompt types are also discussed as an emergent issue that learners raised. These questionnaires are based entirely on self-report.

5.6.1.1 Learners' Perception of the Usefulness and Adequacy of Planning Time

Regardless of proficiency level, learners found planning time useful to task performance. Almost all advanced learners thought that the planning time provided was adequate, while only slightly more than half of the intermediate learners agreed with this.

5.6.1.2 Learners' Perception of How to Spend Their Planning Time

Regardless of proficiency level, planners concentrated on rehearsal and retrieval operations during pre-task planning. These operations include organizing thoughts (38% of advanced planners; 31% of intermediate planners), solving lexical problems (23% of advanced planners; 25% of intermediate planners), and practicing / rehearsing (15% of advanced planners; 19% of intermediate planners). These results conform to those in Ortega (1999, 2005), who found that the advanced and

low-intermediate Spanish L2 learners identified similar planning strategies and feelings about pre-task planning.

The following question items are answered by both planners and non-planners.

5.6.1.3 Learners' Perception of What Would Help Improve Their Task Performance

Many learners, especially non-planners, thought that the opportunity to plan before the task would improve their performance (29% of advanced learners; 57% of intermediate learners). This is especially true for intermediate learners, whose mean is double that of advanced learners. 21% of advanced learners also thought that having their notes or the picture prompts in front of them when they perform the task would be helpful. Group discussion before re-telling the story was also considered to help with task performance by 16% of the advanced learners and 10% of intermediate learners.

4% of advanced learners and 10% of the intermediate learners thought that having more dialogue in the picture prompts would help, while another 4% of advanced learners and 10% of intermediate learners thought that being nervous when they re-told the story prevented them from performing well.

Learners' perception of what would help improve their performance reflects the difficulties they encountered while planning or re-telling the story. Intermediate learners badly needed the opportunity to plan before the task, in order to compensate for the limitations in proficiency. This also indicates a difficulty in attending to the details of the story *and* monitoring their grammar on-line at the same time. The

burden on the working memory is also reflected in learners' preference for the "here-and-now" dimension, i.e. having their notes or picture prompts in front of them when they narrate. This might even imply that even with pre-task planning, planners found it hard to transfer their planning to on-line performance. There is also the possibility of poverty of planning conditions, including inadequate planning time, the lack of vocabulary provided, the lack of the opportunity to discuss in groups. It is also not surprising to find that learners expressed pressure with the opportunity to plan before the task. Learners in this study also report similar listener strategies, such as findings ways to present effectively and to attract the listener. This is again similar to findings in Ortega (1999, 2005).

Comprehension and vocabulary problems are also implied in learners' wish for group discussion and more dialogue in the prompts. Advanced learners also expressed their awareness of their own limitations in English, while both advanced and intermediate learners were aware that they have to deal with their own emotional problems while narrating. This also implies that while learners were conscious of their own errors, they might also be aware of utilizing listener strategies to capture attention, including organizing the story in a comprehensible way, simplifying the story and even using more simple language when narrating.

5.6.1.4 Learners' Perception of the Benefits of Different Prompt Types

From the learners' point of view, using video snippets as prompts for narratives

is much more beneficial and enjoyable than picture sequences. Learners think that the dialogue in *The Simpsons* provided them with lots of vocabulary that could be recycled in their story re-telling (23% of advanced learners; 33% of intermediate learners), and that the language of *The Simpsons* is simple and thus enhanced comprehension (6% of advanced learners; 33% of intermediate learners). Surprisingly, only a small proportion of advanced learners found picture prompts beneficial in that it had an appropriate length (9%), room for imagination (6%) and a clear organization (6%).

From the learners' preference of the different prompt types, the drawbacks of the respective prompts can also be concluded. While videos contained lots of audio-visual stimulation, it could provide too much information for learners to remember, especially compared with picture prompts. Not being allowed to refer to their notes when re-telling the story added burden to their working memory. Moreover, the colloquial American cartoon tested the listening skills of the learners. This may also inhibit their comprehension of the story, especially for intermediate learners.

There are many more drawbacks for picture prompts. Besides the obvious lack of stimulation, the picture sequence does not have lots of dialogue, nor does it provide a lot of phrases or vocabulary for learners to recycle during their speech, thus creating a problem for intermediate learners who do not possess a rich lexical source. The limitation of eight pictures in conveying a complicated message is also felt by the learners, as they showed better comprehension for videos, which provide lots of

information, rather than pictures. The story of the picture sequence might not be clear to many learners.

Generally speaking, learners find that they benefit from video snippets more than picture prompts in re-telling their narratives. There is also a higher mean for task enjoyment with video prompts (videos: 4.13; pictures: 2.93).

5.7 Summary of Questionnaire Data

After examining both the language production and questionnaire responses, In addition to this, learners in this study also conform to the categories identified by Ortega (1999, 2005) in her study on the benefits and limitations of pre-task planning, and also on listener strategies. On the other hand, learners' perception of task difficulty shows that pre-task planning lessens the cognitive demand of the task, whereas enjoying the task also makes task performance less difficult. Video snippets are found to be more enjoyable than picture prompts, although no significance is found for task difficulty.

The following section summarizes the findings on this study once again, and points out the limitations of the present study. The significance of this study for both language testing literature and pedagogical literature is also suggested, and directions given for future research in similar areas.

CHAPTER SIX

CONCLUSION

6.1 Introduction

Pre-task planning is an area which has been heavily discussed in a number of studies (Bygate, 1996; Crookes, 1989; Elder et al., 2002; Ellis, 1987; Foster & Skehan, 1996, 1997; Iwashita et al., 2001; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997, 1999; Tavakoli & Skehan, 2005; Wendel, 1997; Wigglesworth, 1997; Yuan & Ellis, 2003). These studies have shown that pre-task planning has positive effects on L2 learners' language production for the two aspects of fluency and complexity.

Accuracy is not an area which all researchers agree on, but nevertheless, some researchers found that accuracy benefits from pre-task planning under certain conditions (Foster & Skehan, 1996, 1997; Mehnert, 1998; Skehan & Foster, 1997; Tavakoli & Skehan, 2005). These included the type of planning, the type of task, the grammatical feature explored, task complexity, task structure, the length of planning time, and learner proficiency.

There is also the important issue of the trade-off effect between the three key aspects, fluency, complexity and accuracy. Crookes (1989), Foster and Skehan (1996), Skehan and Foster (1997) and Mehnert (1998) claimed that the trade-off effect occurs between complexity and accuracy. However, studies which explored the difference between pre-task planning and on-line planning (Wendel, 1997, as cited in Yuan &

Ellis, 2003; Yuan & Ellis, 2003) suggested an alternative trade-off effect between fluency and accuracy. Their studies have also shown that accuracy is most affected when learners monitor their language production on-line. Since studies to date are heavily focused on learner planning before the task, the obvious next step is to research on on-line planning and its effect on the different aspects of language production.

The need for more research how much learners monitor their speech production during task performance is emphasized once again in the results of this study. This study showed that pre-task planning has positive effects on some measures of fluency (i.e. speed and filled pauses). Learners produce more silence and repair disfluencies when planning is allowed, which contradicts the hypotheses of this study. Concerning learners' language production in the area of complexity, instead of becoming more complex with the opportunity to plan, their language production becomes more complex in cases where *no* planning is allowed, which contradicts the prediction of this study. Last of all, learners' speech does become more accurate with the opportunity to plan before the task, though this is not statistically significant.

6.2 Integrating the Present Study with Previous Studies

This study investigated university Chinese learners in a foreign language learning context. Since Chinese learners of English do not represent L2 learners of other nationalities, the extent to which the results of this study can be generalized to

other learners has to be carefully considered. Nevertheless, this study cast some light on several important issues relating to the role of pre-task planning in task-based performance.

6.2.1 Previous Findings on Fluency

The first issue concerns the operationalization of the three key aspects of language production: fluency, accuracy and complexity, and how they are measured. Although the traditional claim supports the positive effect of pre-task planning on fluency and complexity, these studies utilized different measures. Concerning fluency, Foster and Skehan (1996) reported that planners pause less frequently and spent less time in total silence than non-planners. Skehan and Foster (1997) found the same for total pauses. Wendel (1997) found that planners produce more syllables per minute, and also show a lower mean length of pauses than non-planners. Ortega (1999) found a faster speech rate in planners.

These studies showed that pre-task planning does not have a positive effect on *every* measure of fluency; the main effects are on speech rate and pause length. In addition to this, Tavakoli and Skehan (2005) found that not every measure of fluency is similarly influenced by pre-task planning. They found that pre-task planners are more likely to engage in modification of speech on-line, which is also found in this study. There seems to be a separation between repair fluency and breakdown fluency/speed on the other, suggesting that fluency should be operationalized

carefully.

6.2.2 Present Findings on Fluency

This study also found that fluency is enhanced by pre-task planning, and had an insignificant positive effect on accuracy. In the previous sections, Robinson's (2001) suggested framework of tasks that is different from Skehan's (1998) cognitive approach is mentioned.

According to Robinson (2001), learners have the capacity to attend to more than one aspect of language at the same time. By manipulating "resource-directing" and "resource-depleting" factors, as he called them, the cognitive demand required for task performance will vary, which in turn leads to the variation in quality of the language produced (see Discussion). The opportunity for pre-task planning is counted as a "resource-depleting" factor, which promotes fluency at the expense of accuracy and complexity, while the absence of pre-task planning time, on the other hand, enhances complexity and accuracy but not fluency. The claim that learners will "rise to the challenge" is not supported by the majority of studies on pre-task planning, which found that pre-task planning supports fluency and complexity instead

However, the results of this study do not give enough evidence to support Robinson's (2001) claim, which is contrasted by so many previous studies. Thus more research which focused on task difficulty should be considered before any conclusion can be drawn.

6.2.3 Previous Findings on Complexity

Similarly, researchers studied various measures for complexity. Crookes (1989) measured syntactic complexity and lexical variety. Skehan and Foster (1997) examined the index of subordination. Ortega (1999) explored the number of words per utterance. Yuan and Ellis (2003) studied syntactic complexity and syntactic variety.

6.2.4 Present Findings on Complexity

This study found the lack of pre-task planning effects on complexity in this study. Again, further research in this area has to be conducted before any conclusions can be drawn on concrete issues of pre-task planning effects on complexity.

It can be generalized that not every single fluency or complexity measure is enhanced by pre-task planning *in the same study*. Following Skehan and Foster (1997, 1999) and Tavakoli and Skehan (2005), this study made use of specific measures for fluency (see Method) and generalized measures for complexity. Nevertheless, the amount of research in pre-task planning is huge; for each study to be comparable with the others, there must first be a standardization of the measures used. Before then, conclusions drawn from these studies may possibly be less information than they could be.

6.2.5 Previous Findings on Accuracy

Accuracy is a more complicated issue, as previous research has produced mixed results this measure. A number of studies have reported that pre-task planning has a positive effect on the accuracy of learners' language production (Ellis, 1987; Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997), while other researchers showed that accuracy is not affected by pre-task planning (Skehan & Foster, 2001; Wendel, 1997, as cited in Yuan & Ellis, 2003; Wigglesworth, 1997; Yuan & Ellis, 2003).

From studies which found that accuracy is enhanced with pre-task planning, several generalizations can be drawn as to what variables affect accuracy. Some of these include the type of task (personal or narrative tasks favour accuracy), task complexity (a task which is cognitively less demanding favours accuracy), task structure (structured tasks favours accuracy), the length of planning time (limited planning time leads to increased accuracy), and learner proficiency (advanced learners are able to benefit more from planning).

6.2.6 Present Findings on Accuracy

The present study operationalized some of the variables mentioned above in favour of accuracy, which means that it is highly predicted that an increase in the accuracy of learners' language production would occur. In this study, both advanced and intermediate learners performed a series of narrative tasks with picture sequences

and video snippets which contained an inherent structure in the storyline. They were also given limited planning time for certain tasks. According to the generalizations above, advanced learners performing their tasks with planning time should be able to benefit in the domain of accuracy.

However, results showed that pre-task planning has only a minor effect on accuracy, with an insignificant difference between planners and non-planners, or advanced and intermediate learners. So is accuracy an unstable factor from which we cannot guarantee any enhancement, even with the right conditions?

One of the explanations is the fact that pre-task planners tend to complexify the task, thus leaving less attentional resources for grammatical accuracy (Skehan and Foster 1996). However, a more interesting explanation is offered by Hulstijn and Hulstijn (1984), Wendel (1997) and Yuan and Ellis (2003), who claimed that a key factor to the enhancement of accuracy may be whether learners have the opportunity to plan on-line. Yuan and Ellis (2003) defines on-line planning as follows: "On-line planning is the process by which speakers attend carefully to the formulation stage during speech planning and engage in pre-production and post-production monitoring of their speech acts" (p. 6).

This definition refers directly to Levelt's (1989) model of speech processing (see Literature Review). Allowing learners more time to complete a task is assumed to assist the planning and production of speech, by allowing the search for grammatical linguistic resources, facilitating the process of pre-production monitoring, and

encouraging the process of post-production monitoring (Yuan & Ellis, 2003). Thus language produced in this way will display greater complexity and accuracy, but not fluency.

Yuan and Ellis (2003) suggested that the major difference between pre-task planning and on-line planning is that for the former, learners focus on the first stage of Levelt's model – conceptualization – but may not be able to remember pre-planned forms when they are performing the task. This is a point that the learners of this study reflected upon when they filled in post-task difficulty questionnaires; a small proportion of learners indicate a failure to transfer their planning to on-line performance.

In contrast, on-line planning allows for learners to attend closely to formulation, thus enabling grammatical encodings even during task performance. In short, pre-task planning benefits fluency and complexity, to the detriment of accuracy. On the other hand, on-line planning influences complexity and accuracy positively, but not fluency. As anticipated, Yuan and Ellis (2003) found that the on-line planning group achieved significantly greater accuracy than both pre-task planning and non-planning groups in their study. As the number of studies that explore on-line planning is scarce, this is the obvious next step for further studies, instead of solely focusing on pre-task planning.

6.3 The Trade-Off Effect

6.3.1 Previous Findings on the Trade-Off Effect

An important issue concerns the trade-off effect between the key aspects of language production. There is general agreement that learners with limited L2 proficiency have to decide which aspect of language to prioritize with their limited attentional resources. As mentioned in the previous sections, the trade-off effect is found to occur between accuracy and complexity by many studies (Crookes, 1989; Mehnert, 1998; Skehan & Foster, 1997; Tavakoli & Skehan, 2005), while other researchers claim that fluency and accuracy are traded-off with each other (Wendel, 1997, as cited in Yuan & Ellis, 2003). Yuan and Ellis (2003) even found a dual trade-off effect, the first primarily involving fluency and accuracy, as suggested by Wendel (1997). The second conforms to the mainstream view that accuracy and complexity are in mutual tension with each other.

6.3.2 Present Findings on the Trade-Off Effect

This study found results that supported the traditional claim of accuracy versus complexity. This can be understood as reflecting a basic distinction between a concern whether learners should focus on form (grammatical accuracy) or meaning (fluency and complexity) when they plan (Van Patten 1990, 1996).

To relieve some of the tension between the competing goals and achieve a balance between them, Doughty (1990) and Willis (1996) suggested pre-task activities which are consciousness-raising and which make target grammatical forms salient. Swain (1996) suggested collaborative dialogue during task performance, while

post-task activities which are consciousness-raising, reflective, and analysis-oriented are suggested by Willis (1996), Willis and Willis (1987) and Skehan and Foster (1997) respectively. Yuan and Ellis (2003) also suggested providing learners with both the opportunity to plan *before* and *during* the task. Clearly these guidelines remain to be demonstrated in future research.

6.4 Limitations of the Present Study

6.4.1 The Use of Different Prompt Types

Despite the best efforts, there are several limitations that might be additional factors to the effects of pre-task planning on fluency, accuracy and complexity. This study follows the trend of research on pre-task planning and utilized narrative tasks for speech elicitation. The traditional use of prompts in these narrative tasks are picture sequences extracted from guided composition books (Hill 1960; Heaton 1975). In this study, both picture sequences and video snippets are used as prompts for narrative tasks. The quantitative results obtained from this study shows that the different prompts might benefit different aspects of language production.

For almost every measure of fluency (except for speed), learners performing tasks with video snippets actually conform to the predictions of this study (i.e. non-planners produce more pauses and repair disfluencies than planners). However, these learners produce less complex and less accurate speech when compared with those performing tasks with picture sequences. This suggests that picture sequences

might only be beneficial to certain aspects of language production, but not all.

However, it must be noted that it is difficult to achieve equivalence across the different prompts and planning conditions. As mentioned in the previous section, tasks that deal with picture sequences provide exactly 0.5 minutes and 3.5 minutes for non-planners and planners to look at the pictures and plan their speech respectively, while tasks dealing with video snippets provide an additional 7 minutes or so for both non-planners and planners to watch the video, during which they are also allowed to take notes and plan.

6.4.2 Interpretation of Learners' Perception of Tasks

From the learners' response in the task difficulty questionnaires, videos are definitely the more popular choice in terms of task enjoyment and the benefits (comprehension and vocabulary) they conceived to obtain from the task. In the opposite, picture prompts are criticized as lacking in audio-visual stimulation, vocabulary and is hard to comprehend. This is especially reflective of intermediate learners. If learner attitude is significantly related to learner performance, then the use of video snippets as an alternative prompt for narrative tasks would be worth investigating. Because of this, in addition with the difficulty to operationalize the different prompts, and the scarcity of pre-task planning studies that utilized video snippets as the major tool for testing, further research is needed to support the use of video snippets as a narrative prompt.

Research shows task performance difference according to language expertise (Ortega 1995, 1999). Indeed, intermediate learners show a preference for more enjoyable and less difficult tasks / prompts. Other than this, this study found little difference in the task performance of learners belonging to different proficiency levels, in relation to the benefits of pre-task planning and task performance.

However, learners did express similar pre-task planning strategies as reported by Ortega (1999), and moreover reflected that pre-task planning benefits task performance, and that performing a task which contains a more enjoyable prompt (i.e. videos) lessens task difficulty. Whether this affects their actual task performance or not, is not totally clear in the findings of this study.

It should be noted, however, that learner reactions should be interpreted with caution. Elder et al. (2002) found that the test-taker's perception of task difficulty does not correspond to the hypothesized difficulty of different task conditions, and suggests that we should not rely too heavily on test-taker feedback, "either as a basis for test design or in mounting test validation arguments" (Elder et al. 2002, p. 363). An explanation for this situation is that learners might not all have noticed the difficulty factors, and thus there is an on-going difficulty of making *a priori* estimates of task difficulty.

6.4.3 Pedagogical Studies vs. Language Testing Studies

The difference between the results of this study and that of previous research can

also be understood as a result of differences between testing and pedagogic contexts. Tavakoli and Skehan (2005) concluded that two general problems remain in examining pre-task planning. One of these is the Observer's Paradox, which indicates that there are inconsistencies in the findings between different studies in the literature because of the two sorts of research that can be identified: (1) studies which are pedagogy or acquisition oriented; and (2) studies which replicate testing conditions.

They argued that these two sorts of studies will potentially lead to four relevant sources of differences, which are namely "that testing-linked research leads to a different type of language use; that different types of experimental variables are researched in the two contexts; that common variables have been operationalized differently in the two contexts; [and] that scoring procedures differ in the two contexts" (Tavakoli & Skehan, 2005, p. 33-34).

This study clearly belongs to one which replicates testing conditions which, in the words of Elder et al. (2002), "produce(s) a cognitive focus on display rather than on task fulfillment or getting the message across" (Elder et al. 2002, p. 362).

Under these conditions, learners might be therefore unwilling to exploit possibilities offered by varying task conditions, and instead focus on getting the task done. Because of the different foci, motivations and operationalizations for experimental variables between studies focusing on teaching and testing, more research is needed to distinguish more clearly between these two types of study and standardize the operationalization of variables, before any conclusion can be drawn on

task conditions.

6.5 Implications for Language Pedagogy

Finally, the implications of the results of this study for language pedagogy should be considered. Most researchers agree that giving L2 learners time to plan the task is no doubt beneficial to language production and task performance; however, different types of planning and activities should also be considered. It appears that pre-task planning benefits fluency and complexity, while on-line planning affects accuracy and complexity positively. A variety of pre-task and post-task activities to balance form and meaning, as suggested by Willis (1997), Swain (1998) and Skehan and Foster (1997), should also be integrated into the tasks. However, there is also the argument on "situational authenticity" (Bachman & Palmer, 1996, p. 102), which indicates that learners should be required to engage with tasks under the same conditions that they experience in the real world.

On the other hand, Widdowson (1984) argues that the goal of language teaching should not be to teach language *as* communication, but *for* communication. Thus teachers should aim at developing learners' general linguistic capacity until a balanced goal development is reached (Skehan, 1998). In addition to this, the needs of advanced and intermediate learners should also be taken into account when designing tasks for the classroom. Alternative teaching tools for story re-telling activities, such as videos, should be considered for learners with limited proficiency.

Last of all, learner attitudes may more or less affect language performance. Although the perception of task difficulty is a multi-dimensional phenomenon which results from complex and unstable instructions between different task features and different test-taker attributes (Elder et al., 2002), nevertheless, the qualitative area of learners' perception of task difficulty, and their different usage of pre-task planning strategies, remain a resourceful area for exploration in further studies.

REFERENCES

Anderson, J.R. (1995). *Learning and memory: An integrated approach.* New York: Wiley.

Bachman, L.F., & Palmer, A. (1996). *Language testing in practice.* Oxford: Oxford University Press.

Baddeley, A., & Hitch, G. (1974). Working memory. In G.. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8). New York: Academic Press.

Baddeley, A., & Logie, R. (1999). The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 28-62). New York: Cambridge University Press.

Bradshaw, J. (1990). Test-takers' reactions to a placement test. *Language Testing, 7,* 13-30.

Brown, A. (1993). *The role of test-taker feedback in the test development process: Test-takers' reactions to a tape-mediated test of proficiency in spoken Japanese.* Language Testing Research Colloquium, Tsukuba, Japan, July.

Bygate, M. (1996). Effects of task repetition: Appraising the developing language learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136-146). London: Heinemann.

Castellaneta, D., & Beer, D. (2000). Hell Troupe. In Brooks, J. L. *The simpsons: Too hot for TV.* New York: 20th Century Fox Home Entertainment.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition, 11,* 367-383.

Doughty, C. (1991). Second language instruction does make a difference: evidence form an empirical study on SL relativizaton. *Studies in Second Language Acquisition, 13,* 431-69.

Elder, C., Iwashita, N., & McNamara, T. (2002). Estimating the difficulty of oral

proficiency tasks: What does the test-taker have to offer? *Language Testing, 19*, 347-368.

Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition, 9,* 12-20.

Ellis, R. (2005). *Planning and task performance in a second language.* Amsterdam; Philadelphia: John Benjamins Company.

Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition, 18,* 299-324.

Foster, P., & Skehan, P. (1997). Modifying the task: The effects of surprise, time and planning type on task based foreign language instruction. *Thames Valley Working Papers in English Language Teaching, 4,* 86-109.

Foster, P., & Skehan, P. (1999). The influence of planning on performance in task-based performance. *Language Teaching Research, 3*, 215-247.

Foster, P., & Tonkyn, A. (1997, October). *Measuring spoken language: A unit for all reasons.* Paper presented at the Joint CALR/CALS Symposium, University of Reading, UK.

Gass, S., & Mackey, A. (1999). The effects of task repetition on linguistic output. *Language Learning, 49,* 549-581.

Heaton, J.B. (1975). *Composition Through Pictures.* UK: Longman.

Huizenga, J. (2000). *Can you believe it?: stories and idioms from real life.* New York: Oxford University Press.

Hulstjin, J., & Hulstjin, W. (1984). Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning, 34,* 23-43.

Iwashita, N., & Elder, C. (1997). Expert feedback? Assessing the role of test-taker reactions to a proficiency test for teachers of Japanese. *Melbourne Papers in Language Testing, 6,* 53-67.

Iwashita, N., McNamara, T. & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design. *Language Learning, 21,* 401-36.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: Massachusetts Institute of Technology Press.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition, 20,* 83-108.

Mok, J. (2005). *Predicting task difficulty in speech performance: How do planning time and video prompts influence narrative speech?* Unpublished paper.

Newell, A., & Simon, H. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice Hall.

Ortega, L. (1995). *The effect of planning in L2 Spanish narratives.* Research Note 15. Honolulu HI: University of Hawaii, Second Language Teaching and Curriculum Centre.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21,* 109-148.

Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 77-100). Amsterdam; Philadelphia: John Benjamins Company.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning, 45,* 99-140.

Robinson, P. (2001). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics, 21,* 27-57.

Shohamy, E. (1982). Affective considerations in language testing. *The Modern Language Journal, 66,* 13-17.

Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp.

17-30). London: Heinemann.

Skehan, P. (1998). *A cognitive approach to language learning.* Oxford: Oxford University Press.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Testing Research, 1,* 185-211.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49,* 93-120.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183-205). Cambridge: Cambridge University Press.

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239-276). Amsterdam; Philadelphia: John Benjamins Company.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty and J. Williams (Eds.), *Focus on form in second language acquisition* (p.64-81). New York: Cambridge University Press.

Van Patten, B. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 12,* 287-301.

Van Patten, B. (1996). *Input processing and grammar instruction: Theory and research.* Norwood, NJ: Ablex.

Widdowson, H. (1984). *Explorations in applied linguistics.* Vol. 2. Oxford: Oxford University Press.

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing, 14,* 85-106.

Wigglesworth, G. (2001). Influences on performance in task-based oral assessments.

In Bygate, M., Skehan, P., & Swain, M. (Eds.), *Researching pedagogic tasks : second language learning, teaching and testing* (pp. 186-209). New York: Longman.

Willis, J. (1996). *A framework for task-based learning.* London: Longman.

Willis, D., & Willis, J. (1987). *COBUILD* Book 1, London: HarperCollins.

Willis, D., & Willis, J. (1996). Conscious-raising activities in the language classroom. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 5-33). London: Heinemann.

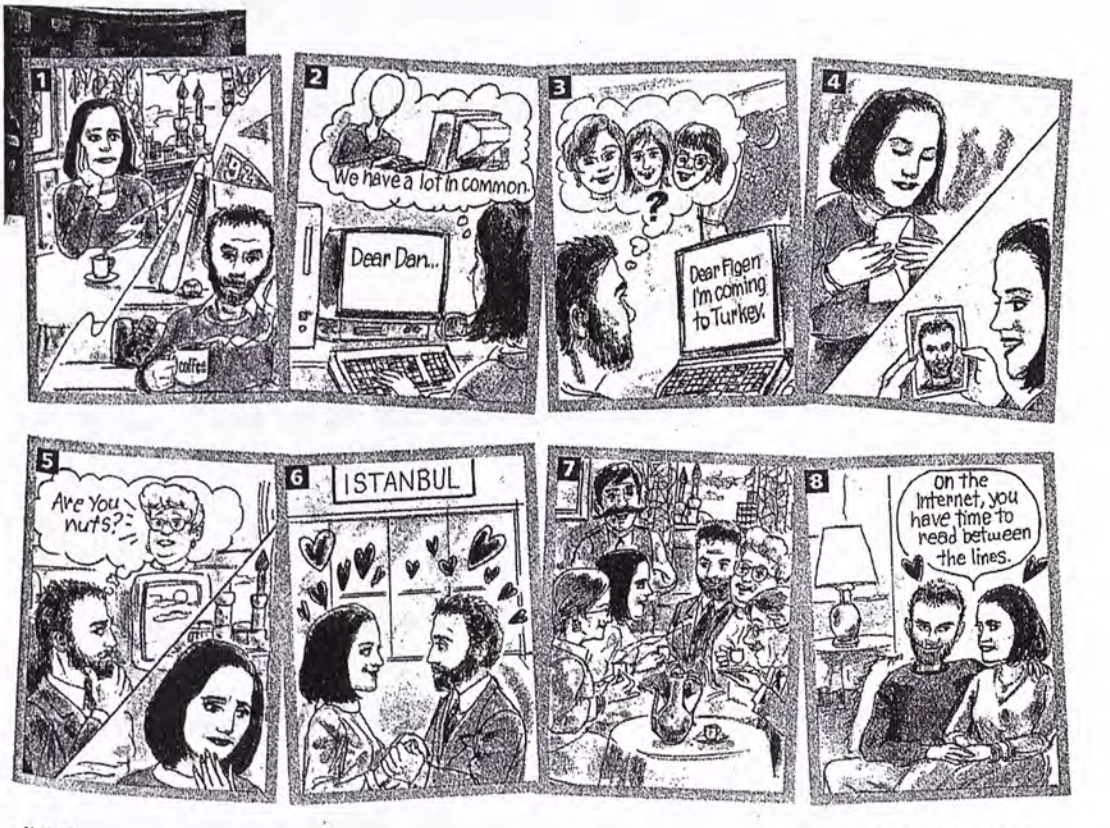
Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24,* 1-27.

Zeidner, M. (1988). Sociocultural differences in examinees' attitudes toward scholastic ability exams. *Journal of Educational Research, 80,* 258-352.

Zeidner, M. (1990). College students' reactions towards key facets of classroom testing. *Assessment and Evaluation in Higher Education, 15,* 151-69.

APPENDIX A

Picture sequence extracted from Huizenga (2000)



Unit 13

Can You Believe It? **81**

APPENDIX B

Instructions given to the participants (English version)

This activity takes around 30 minutes. Your task is to look at the narrative prompt and re-tell the story to me in English. You will be given 2 speaking tasks. One of these tasks is based on a picture sequence, while the other is based on a video clip. You will be given either 0.5 minutes or 3.5 minutes to think about what you are going to say. You have around 3 minutes to tell me the story. Start the story with "this is a story about …" and use the third person perspective in the story. Your speech will be recorded.

After each task, you have to complete a questionnaire on your opinion of the task, the planning time you are given and the picture sequence / video clip you watched. You can now ask me any questions you don't understand.

APPENDIX C

Language Background Questionnaire

The following information will remain confidential.

Name: _____

1. Name of university you currently attend: _____

Major/Year: _____

2. Please list your grades for the following:

A-level Use of English Exam: (Overall) _____

(Oral) _____

3. When did you start learning English? _____

4. Did you attend a CMI or EMI school? _____

5. What courses did you take this semester that involves participation in English?

_____

_____

_____

6. How much are you exposed to English outside classroom? Please list occasions for reading, writing, listening or speaking.

_____

_____

_____

7. Have you traveled abroad to English-speaking countries? Please list countries if you have.

_____

_____

_____

8. How do you feel about using English? (Please circle as appropriate)

A) Inside classrooms (formal setting):

Least comfortable *Very comfortable*

1 2 3 4 5

B) Outside classrooms (informal setting):

Least comfortable *Very comfortable*

1 2 3 4 5

APPENDIX D

Task Difficulty Questionnaire

Name: _____

Group: _____

For planners:

1. Is the planning time useful? (please circle as appropriate) Yes No

2. Is it adequate? (please circle as appropriate) Yes No

3. How did you spend your planning time? Did you think about grammar? Vocabulary? The best way to organize your story? Give examples.

_____

_____

_____

For all participants:

4. How difficult is the task? (please circle as appropriate)

Least difficult *Most difficult*

1 2 3 4 5

5. Did you enjoy the story? (please circle as appropriate)

Least enjoyable *Most enjoyable*

1 2 3 4 5

6. What would help you achieve better performance? You can target at the difficulties that you encountered when telling the story. Did you experience any difficulty in grammar? Vocabulary? Or how to organize your story?

_____

_____

_____

7. Do you have other comments?

_____

_____

_____

APPENDIX E

Consent Form

I understand that MOK Joyce Mee Luen is conducting this study as part of the program requirement for the M.Phil Program in English (Applied English Linguistics). I understand that an MP3 player is used for recording in this study, and that the data collected is solely used for research purposes.

Name _____

Signature _____

Date _____